

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2003, 12:21:43 ; Search time 4228 seconds  
(without alignments)  
16632.840 Million cell updates/sec

Title: US-10-082-894-1  
Perfect score: 1719  
Sequence: 1 gtttaattaccgaagttaga.....aaatctccgtataaaaaatt 1719

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues  
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rnd.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	360.8	21.0	1563	8	AB033336	AB033336 Aspergill
C 2	336.2	19.6	299850	1	AF004601	AF004601 Oceanobac
C 3	321.8	18.7	36296	3	AF039713	AF039713 Caenorhab
C 4	314.2	18.3	248050	1	AL596172	AL596172 Listeria
C 5	314.2	18.3	349980	6	AX413018	AX413018 Sequence
C 6	314.2	18.3	349980	6	AX417048	AX417048 Sequence
C 7	312.6	18.2	324050	1	AL591983	AL591983 Listeria
C 8	312.6	18.2	349980	6	AX641672	AX641672 Sequence
C 9	301.2	17.5	301150	1	AF003190	AF003190 Clostridi
C 10	299.6	17.4	1654	1	AF120090	AF120090 Bacillus
C 11	299	17.4	300854	1	AE017014	AE017014 Bacillus
C 12	292.6	17.0	291205	1	AE017040	AE017040 Bacillus
C 13	291.4	17.0	5111	8	ASPLAS	X64705 Antithamnio
C 14	287.2	16.7	12062	1	AE007586	AE007586 Clostridi
C 15	284.8	16.6	12844	1	U39725	U39725 Mycoplasma
C 16	284.8	16.6	80073	6	AR300198	Continuation (6 of
C 17	283.8	16.5	303650	1	AP001519	AP001519 Bacillus
C 18	282.6	16.4	300050	1	AP004171	AP004171 Mycoplasma
C 19	272.4	15.8	296650	1	AP005082	AP005082 Vibrio pa
C 20	271.2	15.8	300242	1	AE016745	AE016745 Staphyloc
C 21	271	15.8	303671	1	AE015937	AE015937 Clostridi
C 22	269.6	15.7	3373	1	AF269838	AF269838 Staphyloc
C 23	269.6	15.7	3373	6	AX145156	AX145156 Sequence
C 24	269.2	15.7	191028	8	PPU38804	U38804 Porphyra pu
C 25	266.2	15.5	1518	6	AX143497	AX143497 Sequence
C 26	266	15.5	300169	1	AE016801	AE016801 Vibrio vu
C 27	261.8	15.2	37669	1	AC074317	AC074317 Staphyloc
C 28	260.2	15.1	299050	1	AF003131	AF003131 Staphyloc
C 29	260.2	15.1	348527	1	AP003360	AP003360 Staphyloc
C 30	258.6	15.0	290150	1	AP004824	AP004824 Staphyloc
C 31	258.2	15.0	1515	6	AX619984	AX619984 Sequence
C 32	255	14.8	11066	1	AE013608	AE013608 Yersinia
C 33	255	14.8	288050	1	AJ414141	AJ414141 Yersinia
C 34	253.6	14.8	209510	1	BS000018	Z99121 Bacillus su
C 35	253.4	14.7	301903	1	AE016969	AE016969 Mycoplasma
C 36	247.6	14.4	10029	1	AE010963	AE010963 Methanosa
C 37	247.2	14.4	3895	1	BACTPIPMEN	L29475 Bacillus su
C 38	247.2	14.4	347550	1	AP003595	AP003595 Nostoc sp
C 39	245.4	14.3	10322	1	AE013316	AE013316 Methanosa
C 40	242.8	14.1	10701	1	AE000021	AE000021 Mycoplasma
C 41	240.2	14.0	301660	1	AE016768	AE016768 Escherich
C 42	239.4	13.9	304230	1	AE016940	AE016940 Bacteroid
C 43	238.2	13.9	245050	1	AL627280	AL627280 Salmonell
C 44	238.2	13.9	300102	1	AE016847	AE016847 Salmonell
C 45	235.6	13.7	1586	1	AF120091	AF120091 Bacillus

ALIGNMENTS

RESULT 1  
AB033336  
LOCUS  
DEFINITION Aspergillus oryzae gpmA mRNA for phosphoglyceromutase, complete cds.  
ACCSSION AB033336  
VERSION AB033336.1 GI:9955874  
KEYWORDS Phosphoglyceromutase.  
SOURCE Aspergillus oryzae  
ORGANISM Aspergillus oryzae  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.  
REFERENCE 1 (bases 1 to 1563)  
AUTHORS Nakajima,K., Kunihiro,S., Sano,M., Eto,S. and Machida,M.  
TITLE Molecular cloning and characterization of glycolytic gene from

JOURNAL	Aspergillus oryzae					
REFERENCE	Published Only in DataBase (2000)					
AUTHORS	2 (bases 1 to 1563)					
TITLE	Nakajima,K., Kunihiro,S., Sano,M., Eto,S. and Machida,M.					
JOURNAL	Direct Submission					
	Submitted (06-OCT-1999) Keilichi Nakajima, National Institute of Bioscience and Human-Technology, Molecular Biology; Higashi 1-1, Tsukuba, Ibaraki 305-8566, Japan (E-mail:nakajima@nibh.go.jp, Tel:81-298-54-6214, Fax:81-298-54-6240)					
COMMENT	sequence updated (27-Oct-1999).					
FEATURES	Location/Qualifiers					
source	1..1563	/organism="Aspergillus oryzae"				
		/mol_type="mRNA"				
		/strain="RIB40"				
gene	1..1563	/db_xref="taxon:5062"				
		/genes="gpmA"				
CDS	1..1563	/gene="gpmA"				
		/codon_start=1				
		/product="phosphoglyceromutase"				
		/protein_id="BAB12237.1"				
		/db_xref="GI:9955875"				
		/translation="MAKVQKVVLVVDGHWAGVAPDPSRKDGDAILAAETPFMSGFPAEA DSTAQGISLSEKSLAVGLPEGLMGNSEVGHNLINICAGRVWODSVRIDOTLKKGELN KDVNVASIKRAKEGRHLHGLVSDGGVSHNSIHLIGLLVKAKMEIPKPVIFHPF DGDRTSPKATYMQOLLQDKTEKIGICEIATVVGGERVKDDDTLFFFNYSRSDVRREITQ EGESDPVKTIINERYEKDETDFELPIIVGGERRVMDRKRDMRVTAMKGVISQ LLAGDRSPKPPDPYPKNKIHTMTQYKTIDTFPVAFPQHMGVLAEWLSKKDVOQC HVAETKYAHVFPPFGNKIEKQPAGEVRDMISPKVATYDLDPKMSAEAVGQKMADCI ABGEKFWMNFAPPDMVGTGKEYIAIOGVAATDKAIGVIYEACKRGQGVLFITADH GNAEMLTEKPTTPKTSHTTNKVPFINANAPEGWSLKEKGGVLGDVAPTFLAAMGIDQP ENSEEMLLVKA"				
BASE COUNT	362 a	439 c	428 g	334 t		
ORIGIN						
Query Match	21.0%; Score 360.8; DB 8; Length 1563;					
Best Local Similarity	57.%; Pred. No. 4.7e-68;					
Matches	726; Conservative	0; Mismatches	517; Indels	15; Gaps	4;	
QY	181	CTTCATGTTGGATTGCCAGAAGCGTTAAATGGGAAAATCTGAACGTTCGACATTTGCACATA	240			
Db	178	CTTGCTGTCCGATTACCGNAGGCTCTCNGGTTAACAGTGAAGTCGCCCATCTGACACTT	237			
QY	241	GGAGCTGGAGAGATTATTATCAAGATAATGTTTCGAATTAATTTGGCTGTTCACAGAAAC	300			
Db	238	GGTGC GGAGCGTGTGCTGTGCAAGACAGTGTTCGTATCGACCACCCCTCAAGAAGGCG	297			
QY	301	GAGTTTTGTACAATCCTCAGATTGTTGCATCAGCTGACGGTGC AAGAGAGGGAGTGGT	360			
Db	298	GAGTTGAACAAGGTAGACAACTGCTGTTGTCATCCTTCAAGCGTGCCAAGGAAGCAATGCG	357			
QY	361	CGATTCACTTTATTAGCACTGGTTAGCAGTGGTGTCCACCTCTCATATTGATCATCTT	420			
Db	358	CGFTTCATCTCTCGGCCCTGGTCTCCGACGGTGGTGTCCACTCCAACATCACCCACCTT	417			
QY	421	TTTTCGTTGATACGTCGATTTAAACAATTACAAGTCCCAAAGGTTTTTCATTCACTTTTTT	480			
Db	418	ATCGCGCTGTGAAGGTTGCCAAGGAGATGGAGATTCCCAAGGTTTTTCATCCACTTCTTC	477			
QY	481	GCTGATGGTCGAGATACTTCGCCACAAGTGGAGCTGGTTATCTTGAACAACATCTTCAA	540			
Db	478	GCGGATGGAGCTGTACACCGCAACCAAGCGCTTACCAGTACATGCAGCAGCTTCTCGAT	537			
QY	541	TTTATTGCTTCGAAAAGTACGAGAAATTTGGCTACTTATTACTGGACGTTATTATGCAATG	600			
Db	538	CAGACCAAGAAATCGGCATTGCTGAATCGCACCTGTTGTGGACGCTACTGGGCCATG	597			
QY	601	GATAGGGACAAAGATGGGCGGTATTAAGATGGCTTTATGAGGCAATTTGTGAGGATTT	660			
Db	598	GACCGGACAAGCGCTGGGACCGCTGTCGAAATTTGCCATTGCAAGGCGATTGCTCGGGAGAG	657			

TITLE	Genome sequence of Oceanobacillus ihyensensis isolated from the Iheya Ridge and its unexpected adaptive capabilities to extreme environments	
JOURNAL	Nucleic Acids Res. 30 (18), 3927-3935 (2002)	
MEDLINE	22220767	
REFERENCE	12235376	
AUTHORS	3 (bases 1 to 299850)	
TITLE	Takami,H., Takaki,Y. and Chee,G.	
JOURNAL	Direct Submission	
	Submitted (26-DEC-2001) Hideto Takami, Japan Marine Science and Technology Center, Deep-sea Microorganisms Research Group, 2-15 Natsushima-cho, Yokosuka, Kanagawa 237-0061, Japan	
	(E-mail:takami@jamstec.go.jp/jamstec-e/bio/DEEPSTAR/FResearch.html, URL:http://www.jamstec.go.jp/jamstec-e/bio/DEEPSTAR/FResearch.html, Tel:81-468-67-9643, Fax:81-468-67-9645)	
FEATURES	Location/Qualifiers	
source	1..299850	
	/organism="Oceanobacillus ihyensensis HTE831"	
	/mol_type="genomic DNA"	
	/strain="HTE831"	
	/db_xref="taxon:221109"	
	/note="halotolerant and alkaliphilic"	
gene	95..487	
CDS	/gene="OB2343"	
	95..487	
	/gene="OB2343"	
	/note="CDS_ID OB2343"	
	/codon_start=1	
	/transl_table=11	
	/product="hypothetical conserved protein"	
	/protein_id="BAC14299.1"	
	/db_xref="GI:22778029"	
	/translation="MKTIYSPRCSLITSATPLRASSSPSAETSISAPRGPKDITP MIDPKLCLPLEYTSISESNCPALFTKNVGRACPTPSAVINVWCDRTTHPSTPYEL YDCLLSIITLNSHRPTPIYKNHR"	
gene	complement(106..369)	
CDS	/gene="OB2344"	
	complement(106..369)	
	/gene="OB2344"	
	/note="CDS_ID OB2344"	
	/codon_start=1	
	/transl_table=11	
	/product="PTS system, histidine-containing phosphocarrier protein (Hpr protein)"	
	/protein_id="BAC14300.1"	
	/db_xref="GI:22778030"	
	/translation="MKSQTFTTETGVHARPATLLYNKAGQFSEIEVSYKGQVNL KSMGVSLGIPKGAIEVSADGKDEEDALNGVAEVIKEHLGE"	
gene	590..1297	
CDS	/gene="OB2345"	
	590..1297	
	/gene="OB2345"	
	/EC_number="2.4.2.1"	
	/note="CDS_ID OB2345"	
	/codon_start=1	
	/transl_table=11	
	/product="purine nucleoside phosphorylase"	
	/protein_id="BAC14301.1"	
	/db_xref="GI:22778031"	
	/translation="MSVHIGAKQGEIADKILLPGDPLRAKYAETFLLEPKLYNEVRG MYGTGYKGERVSQGTGMVPSISIVNELIQSDYDKLIRVGTGCAIQKQVNRD VLAQAGATTDSOMRLIFNGIDYAPIADFLKKNAYDAGVDKGLHLRVGNVFTSDTFY RDNAAQVNEILLAKYQVLAIEWESTALYTLAAKYCRQALSVLITVSDHIITGEETAEER QTTENDMMETALDILIK"	
gene	1422..2021	
CDS	/gene="OB2346"	
	1422..2021	
	/gene="OB2346"	
	/note="CDS_ID OB2346"	
	/codon_start=1	
	/transl_table=11	
	/product="biotin biosynthesis"	
	/protein_id="BAC14302.1"	
	/db_xref="GI:22778032"	
	/translation="MQQFRAIHITYTSLFVCLMVIGANIAIWFPLTVPYIGGASVPLS LOTPFAILAGMILKGRIGAFAMIVYILIGLTGLPVFAGLKGGLYPMISPTGGFTISFV FIATIVGMLLENKSVSKDLITVAIIGLMINVTGLVYMYIANNEWLELPISLITAM IGWPFVLYKDSVLTIVAAITMMALQKRVSIPVIONISR"	
gene	2117..2596	
CDS	/gene="OB2347"	
	2117..2596	
	/gene="OB2347"	
	/note="CDS_ID OB2347"	
	/codon_start=1	
	/transl_table=11	
	/product="hypothetical conserved protein"	
	/protein_id="BAC14303.1"	
	/db_xref="GI:22778033"	
	/translation="MTLFSNPLMAALCAIVFAOVIKIPIKLIETKEFOPGLAFSTGG MPSSHAAYTALITTAIGITEGTVSSVFALACIFSVITMFEDASGVRRGAGEAVYNQL IRDFQLITTSKADWNKRGTEKQLQELKHGHQPIEVFFGALTGIIILAFSLAPFSY"	
gene	complement(2702..3375)	
CDS	/gene="OB2348"	
	complement(2702..3375)	
	/gene="OB2348"	
	/note="CDS_ID OB2348"	
	/codon_start=1	
	/transl_table=11	
	/product="hypothetical conserved protein"	
	/protein_id="BAC14304.1"	
	/db_xref="GI:22778034"	
	/translation="MKIKNMKMHVGMLSIICLSVLETSVTNLFIPKRVKANPIELIP SSEGIIIEQKNVSLKPKELNVAKANOTLLSNDIEGSSIDEADLEQFPVHTVMTG YTAGIESTGTAHEPGYGTFSGVQVRDLYSTAADLSVYPITGILFIPDYGCVVA DKGSALTGNKIDLYHTVDDVYEQWKKKEVDVYVIEMGDGTLTETLLELNEDESMQV FROQFN"	
gene	complement(3438..3749)	
CDS	/gene="OB2349"	
	complement(3438..3749)	
	/gene="OB2349"	
	/note="CDS_ID OB2349"	
	/codon_start=1	
	/transl_table=11	
	/product="hypothetical conserved protein"	
	/protein_id="BAC14305.1"	
	/db_xref="GI:22778035"	
	/translation="MVOLIVSVFLYFVFFIGFIAFLNMLRQTLMAEYPIVWILII DNVGMAEYFTNTATAFNDLSTFLDLADIVILAGFIGTIGVSFVTKFLKRSYQM F"	
gene	complement(3849..5060)	
CDS	/gene="OB2350"	
	complement(3849..5060)	
	/gene="OB2350"	
	/note="CDS_ID OB2350"	
	/codon_start=1	
	/transl_table=11	
	/product="NADH dehydrogenase"	
	/protein_id="BAC14306.1"	
	/db_xref="GI:22778036"	
	/translation="MNKPHIVILGAGYGGMAAARMOKLIHSNEAKITLVNKNDYHYO TTWLHENAAGTLHHRSDKIDIKVIDRNRKILIDEVYSIKREKKYKVKKQELYYDI LIVSLGEPALYIGPLUSEFAYSINDINARLIERIEINFALTYENKPDGRLNIV IGGLGLGVFEAGLINRPTICEYDIEKNLVKYLLEGRDCVLYGYDDSAVOYAMN SLQSRGVEIINDATVKAEDSIIEYKGSSEHKISTKSVIWTAGTKANSIVERLNLMS EHGVRKYRNLRSIEDDSIYLIIGDCSLNYDENGVLPSAQIAMQOGEFIARNNAKI KGMKMAFLPDVPGCRITSLGSDDGIAOMFSRNLGKWAALKLRVADNWNLYFLKGGIQL LMKKGKFNMFH"	
gene	5424..6410	
CDS	/gene="OB2351"	
	5424..6410	
	/gene="OB2351"	
	/note="CDS_ID OB2351"	
	/codon_start=1	
	/transl_table=11	
	/product="thioredoxin reductase"	
	/protein_id="BAC14307.1"	

```
/db_xref="GI:22778037"
/translation="MDSKVFDTIIGAGTGLTFAFYGCMREASVKIIESLPHIGSQL
TALYPEKYIYDAGFPKPAQELIDRLEQAKFFDPEIVLGOAIDKVERLEDETIRLT
SNTGEVHLTKTIIITAGNACFPRLNIGNSEAFCKNLHYVVRDMNOYKGRVLLG
GSDSAWALMLEKIAAKVTLVRRDQFRAHESVNLKESSEVILLTPYAPINVEASD
KIEKLOQVEKEKIDIVEDDIIICNGYFISLGPDKMGLEIEKNSIIVNSKMETNI
PGVYAGDVCTYDGRVKVLATGFGEGPTAINNAKNYIDPKARIQPKHSTAMF"
gene
CDS
complement(6465..7061)
/gene="OB2352"
complement(6465..7061)
/gene="OB2352"
/codon_start=1
/transl_table=11
/product="chitooligosaccharide deacetylase (modulation
protein).id="BAC14308.1"
/db_xref="GI:22778038"
/translation="MFQORMIALTDFDGPLGPTNEILDILSENQSRATFVLGN
VTKHRDVRRIVREGHIDIGVSHFSPNFQETIEYVYRELIDSTLLELDLGLVLY
RPPYGFSTAETQFLDQIGLSILWYDTFLDWEILNAEOIASNLLNIQAGDIVLLHD
IHPTIOAMRIAPKLVDQGYKLVTISEWLEQENPFK"
complement(7514..7870)
/gene="OB2353"
complement(7514..7870)
/gene="OB2353"
/codon_start=1
/transl_table=11
/product="hypothetical conserved protein"
/protein_id="BAC14309.1"
/db_xref="GI:22778039"
/translation="MIINTESEQIKEMKDESDQSHLRFKIRGGGSGLSYSLGF
DEDINEELDIVIONINGIPVVFNFQDPIEGTTIDYKENLWGGGFSIDNPNAIVSGC
GSSFRKDKVGPQDC"
complement(7943..8803)
gene
Query Match 19.6%; Score 336.2; DB 1; Length 299850;
Best Local Similarity 54.3%; Pred. No. 4.5e-63;
Matches 752; Conservative 0; Mismatches 618; Indels 15; Gaps 3;
QY 65 TTATTCATGATGGGCGCTTCGGATGAACACACCGGAATCAATGCTAAAGCTAA 124
Db 92365 TTTCTTGATGATTCGCTATTCGTGATGAAGTAAAGGAATGCGATCAACAGCAACA 92306
QY 125 CGCTATTATGGCAAACTTTGCTCGAAATGGCAAA---AATTGGAAGCACACGGTC 181
Db 92305 CCCCTAAATTTGATCGTTTGGAAATCAATATGCTATCAATCAATAGAGGCTTCTGGAA 92246
QY 182 TTCATGTTGGATGCCAGAGGCTTAATGGGAAATTCGAAGTTGGACATTTGAATPATG 241
Db 92245 AGGCTGTTGGATTACAGATGGCAAAATGGGAAATCCGAAGTTGGTCATTTGAATATTG 92186
QY 242 GAGCTGGAAGACTTATTAATCAAGATATGTTGGAATTAATTTGGCTGTTCAACGAAACG 301
Db 92185 GTGCTGGTTCGATTTGTTGATCAAGAGCTTACTCGTGAATATGTTGATGACGAGGTTG 92126
QY 302 AGTTTGTACAAATCCTCAGATGTTGTCATGAGCTGAGCGTCAAGAGGAGGAGTGGTC 361
Db 92125 ATTTCTATGAGTTGATGCGTTTCATTAATCTGTAGAATCGAAGAACATGACAAAG 92066
QY 362 GATTCGATTTATGAGCTGGTTAGCGATGGTGGTGGTTCACCTCATATGATGATCTTTT 421
Db 92065 CGCTTCACCTATTATGTTTATTCATGACGGTGGTGGTATGATGATGATGATGATGAT 92006
QY 422 TTGCGTTGATACGTCATTTAAACAAATTAACAGTCCAAAGGTTTTCATTCCTTTTGG 481
Db 92005 TTGCTTTACTTCGATTAGCGAAGAACATGAATTAAGAAATGATTGTACATGCTTTCT 91946
QY 482 CTGATGGTGGAGATCTTCGCCAACAAAGTGGAGCTGGTTATCTTGAACAACTTCCTCAAT 541
Db 91945 TAGACGGAGCTGATGTTGGTCCGAAACACAGCGCAAAATATATCAAGAAACACAGATG 91886
```

## RESULT 3

AF039713

LOCUS

DEFINITION

AF039713

VERSION

AF039713.1

KEYWORDS

HTG

SOURCE

Caenorhabditis elegans

36296 bp DNA linear INV 26-DEC-2002

complete sequence.

cosmid F57B10,

complete sequence.

cosmid F57B10,

complete sequence.

ORGANISM Caenorhabditis elegans  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;  
 Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.  
 REFERENCE 1 (bases 1 to 36296)  
 AUTHORS Waterston,R.  
 TITLE Genome sequence of the nematode C. elegans: a platform for  
 investigating biology. The C. elegans Sequencing Consortium  
 Science 282 (5396), 2012-2018 (1998)  
 JOURNAL 99069613  
 MEDLINE  
 PUBMED 9851916  
 REFERENCE 2 (bases 1 to 36296)  
 AUTHORS Greco,T., Elliott,G. and Keppeler,D.  
 TITLE The sequence of C. elegans cosmid F57B10  
 JOURNAL Unpublished (2001)  
 REFERENCE 3 (bases 1 to 36296)  
 AUTHORS Waterston,R.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-DEC-1997) Department of Genetics, Washington  
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 REFERENCE 4 (bases 1 to 36296)  
 AUTHORS Waterston,R.  
 TITLE Direct Submission  
 JOURNAL Submitted (19-JUL-2001) Department of Genetics, Washington  
 University, Genome Sequencing Center, 4444 Forest Park Avenue, St.  
 Louis, MO 63110, USA  
 REFERENCE 5 (bases 1 to 36296)  
 AUTHORS Waterston,R.  
 TITLE Direct Submission  
 JOURNAL Submitted (19-APR-2002) Department of Genetics, Washington  
 University, Genome Sequencing Center, 4444 Forest Park Avenue, St.  
 Louis, MO 63110, USA  
 REFERENCE 6 (bases 1 to 36296)  
 AUTHORS Waterston,R.  
 TITLE Direct Submission  
 JOURNAL Submitted (20-JUL-2002) Department of Genetics, Washington  
 University, Genome Sequencing Center, 4444 Forest Park Avenue, St.  
 Louis, MO 63110, USA  
 REFERENCE 7 (bases 1 to 36296)  
 AUTHORS Waterston,R.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-OCT-2002) Department of Genetics, Washington  
 University, Genome Sequencing Center, 4444 Forest Park Avenue, St.  
 Louis, MO 63110, USA  
 REFERENCE 8 (bases 1 to 36296)  
 AUTHORS Waterston,R.  
 TITLE Direct Submission  
 JOURNAL Submitted (19-NOV-2002) Department of Genetics, Washington  
 University, Genome Sequencing Center, 4444 Forest Park Avenue, St.  
 Louis, MO 63110, USA  
 REFERENCE 9 (bases 1 to 36296)  
 AUTHORS Waterston,R.  
 TITLE Direct Submission  
 JOURNAL Submitted (26-DEC-2002) Department of Genetics, Washington  
 University, Genome Sequencing Center, 4444 Forest Park Avenue, St.  
 Louis, MO 63110, USA  
 COMMENT Submitted by:  
 Genome Sequencing Center  
 Department of Genetics, Washington University  
 St. Louis, MO 63110, USA, and  
 Sanger Centre, Hinxton Hall  
 Cambridge CB10 1RQ, England  
 email: rwnematode.wustl.edu and jes@sanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone.  
 It may be shorter because we only sequence overlapping sections  
 once, or longer because we provide a small overlap between  
 neighboring submissions.

This sequence was finished as follows unless otherwise noted: all  
 regions were double stranded, sequenced with an alternate chemistry  
 or covered by high quality data (i.e., phred quality >= 30); an  
 attempt was made to resolve all sequencing problems, such as  
 compressions and repeats; all regions were covered by sequence from

more than one ml3 subclone.

For a graphical representation of this cosmid sequence and its  
 analysis see:  
<http://www.wormbase.org/db/seq/sequence?name=F57B10;class=Sequence>

#### NEIGHBORING COSMID INFORMATION

The 5' cosmid is T10E9, 300 bp overlap; the 3' cosmid is F48A9,  
 4100 bp overlap. Actual start of this cosmid is at base position  
 197 of F57B10; actual end is at 6600 of F48A9.

#### NOTES:

Coding sequences below are the result of integration and manual  
 review of the following data : computer analysis using the program  
 Genefinder (P. Green and L. Hillier, personal communication), the  
 large scale EST projects of Yuji Kohara  
 ([http://www.ddbj.nig.ac.jp/c-elegans/html/CE\\_INDEX.html](http://www.ddbj.nig.ac.jp/c-elegans/html/CE_INDEX.html)) and The C.  
 elegans ORFome cloning project (<http://worfdb.dfci.harvard.edu/>),  
 similarity to other proteins from BlastX analyses  
 (<http://blast.wustl.edu/>), sequence conservation with C. briggsae  
 using Jim Kent's WABA alignment program (Genome Research  
 10:1115-1125, 2000), individual C. elegans GenBank submissions,  
 and personal communications with C. elegans researchers. TRNAS  
 are predicted using the program tRNAscan-SE (Lowe, T.M. and  
 Eddy, S.R., 1997, Nucl. Acids. Res., 25, 955-964).

#### FEATURES

Location/Qualifiers  
 source

1. .36296  
 /organism="Caenorhabditis elegans"  
 /mol\_type="genomic DNA"  
 /strain="Bristol N2"  
 /db\_xref="taxon:6239"  
 /chromosome="I"  
 /clone="F57B10"  
 /complement(417..1966)  
 /gene="F57B10.8"  
 complement(join(417..623,676..757,801..970,1018..1119,  
 1648..1799,1873..1966))  
 /gene="F57B10.8"  
 /standard\_name="F57B10.8"  
 /note="contains similarity to Oryctolagus cuniculus  
 TROPONIN T, CARDIAC MUSCLE ISOFORMS (TNTC)."  
 SW:TRT2.RABIT; coded for by the following C. elegans  
 cDNAs: CEESL47F"

gene

CDS

/codon\_start=1  
 /product="Hypothetical protein F57B10.8"  
 /protein\_id="AAB96725.1"  
 /db\_xref="GI:2773208"  
 /db\_xref="WormBase:F57B10.8"  
 /translation="MAPIPKKKIVKSIKKPVVEEDMKMEKEKLEPKEESIETD  
 NNEDEAGNLDSDSEIKTESEDEDETEDEQKOKTKGVVFFSIPKYNVVRREYF  
 EKRCPGVGRIFLARNKSNPQYLYSEGWMELAKKIAKIAEQIDNTPIGKKNKDP  
 VSNMLNRYLSGFEVHLMQOYKEVEKRMNVEVAQARRIAAHFEQIEKRGKHL  
 RKLKAVTAGGKWDKFDQVQKQKVVKKVKKERSKHVNTESSELMNNIFK"  
 complement(2302..4228)  
 /gene="F57B10.9"  
 complement(join(2302..2636,2689..2796,2917..3285,  
 3332..3536,3616..3711,3756..3837,3943..4067,4118..4228))  
 /gene="F57B10.9"  
 /standard\_name="F57B10.9"  
 /note="contains similarity to Interpro domain IPR000169  
 (Eukaryotic thiol (cysteine) protease); coded for by the  
 following C. elegans cDNAs: yk281e10.5"

gene

CDS

```

ERLIELHIVLREFTDFEQABPSTELSKKEKNRISTTISNFIKGGOKIANGVETT
TVRISVENDGQYRTTLTATDKPMQSPVIGKSVYMHKGTAKTRYLLDKIG
DMGVSKKLIADSAORREGDKSGSLVSGTIEILGGITCVATVMSLEDGSRHLCS
IANQTVQNYKLKYGDDASDTHHALFAAGHGTLAAQAQLMDLGRPSRVAGRWARKAGIQM
VSDIHSRSGSLRDQVEPALEKKAV
/genes="F57B10.14"
/standard_name="F57B10.14"
/notes="coded for by the following C. elegans cDNAs:
yk1069d02.5, yk1069d02.3, yk1107g08.3, yk1019c04.3"
/codon_start=1
/product="Hypothetical protein F57B10.14"
/protein_id="AA081096.1"
/db_xref="GI:21913056"
/translation="MLYMKPAREVKNFTNPVFKSYGAYGGSFAFLIAYFCWEKTVG
QYIPLWNKRYVE"
/complement(5230..5905)
/genes="dad-1"
/complement(join(5230..5360,5695..5905))
/genes="dad-1"
/standard_name="F57B10.10"
/notes="contains similarity to Pfam domain PF02109 (DAD
family); coded for by the following C. elegans cDNAs:
CEESW25F, yk115g3.5, X89080"
/codon_start=1
/product="Dad (defender against apoptotic death) homolog
protein 1"
/protein_id="AAB96727.1"
/db_xref="GI:2773210"
/db_xref="WormBase:F57B10.10"
/translation="MAQVVPVLSKLFDDYOKTSSKLIIDAYMTYIILFTGFIQFIY
CLLVGTFPNSFLSGFISVTSFLASCLRMQVQNENSEFTAVSTERAFADFIANL
ILHLVVVNFILG"
/genes="F57B10.7"
/join(6517..6545,6591..6691,7177..7400,7446..7546,
7593..7685,8307..8720,8904..9207,9252..9542,9801..9947)
/genes="F57B10.7"
/standard_name="F57B10.7"
/notes="contains similarity to Pfam domain PF01204
(trehalase); coded for by the following C. elegans cDNAs:
yk86a7.5, CESE59F, yk195g8.3, yk195g8.5, yk375a12.5,
U51925"
/codon_start=1
/product="Hypothetical protein F57B10.7"
Query Match 18.7%; Score 321.8; DB 3; Length 36296;
Best Local Similarity 65.0%; Pred. No. 8.6e-60;
Matches 530; Conservative 0; Mismatches 242; Indels 43; Gaps 2;
QY 667 AAGCCACCGCTGATAGGCTGTCGATGTTGTTAGAGAGCGATATGCTCAATCTGAGACT 726
DB 22974 AAAATTATCATTTTCCAGGTGTTGAAGTTGTTGCGAAGCGTTACGCTGCTGATGAACA 23033
QY 727 GACGATCTTCTGAACCAATGTTTTCGGAGCAT--GGGCGAGTAAAGATGACGAT 783
DB 23034 GACGAATCTTGAAGCCATCATCTTCGAAGAGAAAGAGCGTGTTCAAATGACGAT 23093
QY 784 ACTCTTTATTTCTCAATATGCTGCTGATCGTCGATGCTCAATTTGTGAATGTTTGGGT 843
DB 23094 ACAATCATCTCTTCGACTACCGTGTGATCGTATCGTGAGATTTCGACGAATGGA 23153
QY 844 CTCGACGTTATGAAGATCTTAATAGTTCGTTCCACCCGTAATAATATTCAGATAGT 903
DB 23154 ATGGATCGTTACAAGGATTCGATTCGAAGTTAGCTCATCCATCAATCTTCAAGTATAT 23213
QY 904 GGGATGACCAATACATAAAGAGTTTCCATTTCCATGCTTATCCACCTGTGACTCAT 963
DB 23214 GGATGACTCATACAAAGCCGAGTTCCCATTCGAATTCGATGTTCCCGCCGACGATGAAC 23273

```

```

QY 964 ACTAATGTGCTGCTGAATGGCTGCTTCTCAAGGAGTACTCAATTTTCACTGTGGGAA 1023
DB 23274 AAAATGATTTGGCTGAGTGGCTCGCGGACCAAAAGTTTCGCAATTTTCACTGTGGGAA 23333
QY 1024 ACTGAGAAGTATCTCTCATGTTACCTTCTTTTAAATGGTGGTGGAGGTTTCAATTCCAA 1083
DB 23334 ACCGAAAAATACGCTCAGCTTACATTTTCTTCAATGGAGGACTTGAAAAACAATTTGAG 23393
QY 1084 GATGAAGAGCGTGTATGTTCCGTCACCAAAAG----- 1117
DB 23394 GGAGAAGAAGGTGTTTGTAGTGCCAGTCCAAAGGTAATATTGATGATATAACTTCGAA 23453
QY 1118 -----AAGTTGCTACATATGATTTTAAACACCAAGAAATGAATCTGCTCGAGT 1163
DB 23454 AATTGAATATTTCCAGGTGCGAATTTACGATCTTCAACCAAGAAATGCTTGGGCGCGCT 23513
QY 1164 TGGCGAAAAAATGGTGCAGCAAAATGAGTCAGGCAGGACATCTTTTGGTTATGTGCAATTT 1223
DB 23514 TGTGTGAAAAATGATTTGAACAACCTCGAGGCTGGAACATCATCTTATTTATGTGCAACTT 23573
QY 1224 TGGGCTCTCTGACATGTTGGACATACCTGTAATTTGAACCTGCGCTCAAGCATGTCA 1283
DB 23574 TGCTCCACGAGATATGTTGGGATACGGGAGTCTATGAAGCTGCTGTTCAAGGCTGTGA 23633
QY 1284 AGCTACTGACGAGGCAATTTGAAAGATATTTGAAGCATGCCAAACTTATAATTAACCTTCT 1343
DB 23634 AGCTACTGATATCGCAATCGGAAGATCTATGAAGCACTCAAAAGCATGATACACT 23693
QY 1344 TATGGTTACTTCGATCATGGAATGCTGAGAAATGATGTTGCTCCGATGGTAGTGAACA 1403
DB 23694 TATGGTTACTGCTGATCAGGAAATGCTGAAAGATGAAGGCTCCAGATGGTGAACA 23753
QY 1404 TACTGCACATACCTGCAATTTGGTCCCAATTTACTT 1438
DB 23754 CACTGCTCACATGTTACCGTGTTCCTCACTCACTT 23788

RESULT 4
AL596172/c 248050 bp DNA linear BCT 06-JUN-2002
LOCUS Listeria innocua CLIP11262 complete genome, segment 10/12.
DEFINITION AL596172 AL592022
ACCESSION AL596172.1 GI:16414852
VERSION
KEYWORDS
SOURCE Listeria innocua
ORGANISM Listeria innocua
REFERENCE 1
AUTHORS Glaser,P., Frangeul,L., Buchrieser,C., Rusniok,C., Amend,A.,
Baquero,F., Berche,P., Bloeker,H., Brandt,P., Chakraborty,T.,
Charbit,A., Chetouani,F., Couve,E., de Daruvar,A., Dehoux,P.,
Domann,E., Dominguez-Bernal,G., Duchaud,E., Duranc,L.,
Dussauguet,O., Entian,K.D., Fsihi,H., Portillo,F.G., Garrido,P.,
Gautier,L., Goebel,W., Gomez-Lopez,N., Hain,T., Hauf,J.,
Jackson,D., Jones,L.M., Kaerst,U., KrefT,J., Kuhn,M., Kunst,F.,
Kurapkat,G., Madueno,E., Maitournam,A., Vicente,J.M., Ng,E.,
Nedjari,H., Nordstiek,G., Novella,S., de Pablos,B., Perez-Diaz,J.C.,
Purcell,R., Rammel,B., Rose,M., Schlueter,T., Simoes,N., and Cossart,P.
Tierrez,A., Vazquez-Boland,J.A., Voss,H., Wehland,J. and Cossart,P.
Comparative genomics of Listeria species
Science 294 (5543), 849-852 (2001)
21537279
PUBMED 11679669
REFERENCE 2 (bases 1 to 248050)
AUTHORS Glaser,P., Frangeul,L. and Rusniok,C.
TITLE Direct Submission
JOURNAL Submitted (09-JUL-2001) Glaser P., Institut Pasteur, Genomique des
Microorganismes Pathogenes, 25 rue du Docteur Roux, 75724 Paris
Cedex 15, FRANCE
COMMENT E-mail: pglaser@pasteur.fr
FEATURES Phone: +33 (0)1 45 68 89 96, Fax: +33 (0)1 45 68 87 86.
Location/Qualifiers
1..248050
source

```



Query Match	18.3%;	Score 314.2;	DB 1;	Length 248050;
Best Local Similarity	53.3%;	Pred. No. 2.8e-58;		
Matches 767;	Conservative 0;	Mismatches 648;	Indels 24;	Gaps 4;
QY	14	AGTTTGAGATGACAAATATCAAAATGTTCAACAAAAGTCTGCTTGTAGTTATTGATG 73		
DB	214453	ATTATTGGAGGCAGTAAATAAATAGTAGTAATCACCTGTAGCAATATTATTCTCGATG 214394		
QY	74	GATGGGGCTTTCCGATGAACAACACGCGAATGCTAAAGCTTAAACACGCCATTATTA 133		
DB	214393	GTTTGGTAAAGCTGCAGAAACAGTAGTAATGCTGTAGCTCAAGCAACAACCAAAATT 214334		
QY	134	TGCACAAACTTTTGTCTGCGAAATTGGCA---AAATTTGGAAGCACACAGGCTTTCATGTTG 190		
DB	214333	TCGACCGTTACTGGGCTAATTTCTCTCACGGGGAACTTAAAGCTGTGGCTTTGATGTTG 214274		
QY	191	GATTCGCCAGAGGCTTAATGGAATTCGAAGTTGGACATTTGCAATATAGGACGTGGAA 250		
DB	214273	GTCTTCCAGAGGCCAAATGGGTAACTCTGAAGTTGGCCATACAACATCTGGAGCTGGAC 214214		
QY	251	GAGTTATTATTCAAGATATTGTTCCGAATTAATTTGGCTGTTCAACGAAACGAGTTTGTTA 310		
DB	214213	GTATTGTCTACCAAGCTTAACCTCGTATTGATAAGCAATTGAAGAAGCGGAATTCCAAG 214154		
QY	311	CAAAATCCTCAGATTGTTCCATCAGCTGAGCGTGCAAAGAAGGGAGTGTGTCGATTCGATT 370		
DB	214153	AGAACAAAGCCCTAAACAATGTTTCACTCATACAAGAAACAAATTCGACCTACATC 214094		
QY	371	TATTAGGACTGGTTAGCGATGGTGGTGCACCTCTCATATTGATCATCTTTTGGCGTTCA 430		
DB	214093	TTTTTCGGTTTACTATCAGACGGCGCGTGCATAGTCACATTAACTACCTCGTTGTCACITTT 214034		
QY	431	TACGTGCATTTAAACAATTAACAAGTGCACAAAGTTTTCATTCCTACTTTTGTGATGGTC 490		
DB	214033	TAGAAACTGCGAAAGATAAAGCGGTGAAAAACGTTTATATCCATGCATCCCTTGATGGAC 213974		
QY	491	GAGATACCTTCGCCAACAGTGGAGCTGGTTATCTTGGAACAACCTTCTCAATTTATGCTTT 550		
DB	213973	GTGACGTGGCACCAATCCCTCCTACAGAAATTTTGAACACCTAGAAAAAGCTATTAGTG 213914		
QY	551	CGGAAAAGTCGAGGAATTGGCTACTATTACTGAGCTTATTATGCAATGGATAGGACA 610		
DB	213913	ACTTAAACTATGGTGCATTTGCTACTGTTCTGACGTTTCTACGGGATGGATCGTGATA 213854		
QY	611	AAAGATGGGACGCTATTAAAGATGGCTTATGAGGCAATGTTGGAGGTATTGGACAAAAAG 670		
DB	213853	AACGCTGGGAACGTTTGAAAAACCATACAAGCAATCTGAACGCTGAAGGTGAAAAA- 213795		
QY	671	CCACGTTTGATAAGGCTGTCGATGTTCTTGAAGAGCGATATGCTCAATCTGAGACTGAGC 730		
DB	213794	-----TTTGAAGATCCCAATCGAATCTGTAAGCTTCTTACGCTAATGACAAAAATGATG 213740		



BASE COUNT	104581 a	69113 c	62414 g	113872 t	
ORIGIN					
Query Match	18.3%;	Score 314.2;	DB 6;	Length 349980;	
Best Local Similarity	53.3%;	Pred. No. 2.7e-58;			
Matches 767;	Conservative 0;	Mismatches 648;	Indels 24;	Gaps 4;	
Qy	14	AGTTTGAGTGGACAAATATCAAAATGTTCAACAAAAGTCTGTCTGTAGTTATTGATG	73		
Db	138870	ATTATTGGAGCAGTAATAATAATGAGTAATACACCTGTAGCAATATTATTCTCTCGATG	138811		
Qy	74	GATGGGCGCTTCCGATGACACACACCGGAATGCAATTTGAAAGCTTAAAGCTTAAACGCGCTATTA	133		
Db	138810	GTTTTGTAAACGTCGAGAACAGTAGTAATGCTGTAGCTCAAGCAACAAACAAAT	138751		
Qy	134	TGGCAAACTTTGTCTCGAAATTTGGCA---AAAATTGGAGACACACGGCTTTCATGTTG	190		
Db	138750	TCGACCGTTACTGGCTAATTTCTCAGGGGAACCTTAAGCTGTGGCTTGATGTTG	138691		
Qy	191	GATTGCCAGAGGCTTAATGGGAAATTTCTGAAGTTGGACATTTTGAATATAGGAGCTGGAA	250		
Db	138690	GTCTTCAGAGGCAAAATGGGTAACTCTGAAGTTGGCCATACAAACATCGGAGCTGGAC	138631		
Qy	251	GAGTTATTATCAGATATGTTTCGAATTAATTTGGCTGTTCACGAAACAGAGTTGTTA	310		
Db	138630	GTATTGTCTACCAAGAGCTTAACCTGATTTGATAAGCAATTTGAAGAAGGCGGAATTC	138571		
Qy	311	CAAACTCTCAGATTTTGCATCAGCTGAGCGTGCAAGAGAGGGAGTGGTCGATTGCATT	370		
Db	138570	AGACAAAGCCCTAAACATGCTTTCATCTACATACAAAGAAACAAATTCGGACCTACATC	138511		
Qy	371	TATTAGGACTGGTTAGCGATGGTGTGCCACTCTCATATTGATCATCTTTTTCGGTTGA	430		
Db	138510	TTTTCGGTTTACTATCAGACGGCGGTGCATAGTACATTAATCACTGTTGCACCTT	138451		
Qy	431	TACGTGATTTAAACATTTACAAGTGCACAAAGTTTTCATTCACCTTTTTCGTGATGGTC	490		
Db	138450	TAGAACTGCGAAGATAAAGGCGGTGAAAACGTTTATATCCATGCAATTCCTTGATGGAC	138391		
Qy	491	GAGTACTTCCGCAACAAAGTGGAGCTGGTATTCTTGAACAACTCTTCAATTTATGCTT	550		
Db	138390	GTGACGTGGCACCAATCTCCTACATAGATAATTTAGAAACACTAGAAAACCTATTAGTG	138331		
Qy	551	CGGAAAAGTACGGAGAAATGGCTACTATTACTGACGCTTATTTATGCAATGATAGGACA	610		
Db	138330	ACTTAAACTATGTGCAATGCTACTGTTCTGCAAGCTTTCTACGCGATGATGCTGATA	138271		
Qy	611	AAAGATGGGACGCTTTAAGATGCTTATGAGGCAATTTGTTGGAGTATTGGACAAAAG	670		
Db	138270	AACGCTGGAAACGTTGTGAAAANGACATACAAAGCAATCGTAAGCGCTGAAGGTGAAAA	138212		
Qy	671	CCACCGTTGATAAGCGTGTGATGTTGTAGAGGCGATATGCTCAATCTGAGACTGACG	730		
Db	138211	-----TTTGAAGATCCCAATCGAATGTTGCAAGCTTCTTACGCTTAATGACAAAATGATG	138157		
Qy	731	AATTTCTGAAACCAATTTGTTTTTCGGACGATGGGC-----GAGTAAAAGATG	778		
Db	138156	AATTCGTTGTCTCTGCTATCATATTACTAAAGATGCAAAACCTGTTGCAACATTTAAAGACA	138097		
Qy	779	ACGATCTCTTATTTTCTTCAATTTATGCTGATGATGATGCGTCAATTTGTTGAAATGTT	838		
Db	138096	ACGATGAGTGAATTTTCTTCAATTTCCGCTCGACCGGATTCACCTTTCTTAATGCAT	138037		
Qy	839	TGGGTCTCGAAGCTTATAAGATCTTAATAGTTCGGTTCCTCACCTTAAAATATTTCAGA	898		
Db	138036	TCACTGATAAAGATGGATCATTTTCGACCGTGGGCTTAACCCCTTAAAACATCAAT	137977		
Qy	899	TTAGTGGGATGACCAACATACAAATAAGAGTTTCCATTTCCATGCTTATTTCCACCTGTGA	958		
Db	137976	TCGTTTACAATGACTCTTTACAAACCAAGCATTTGATGCTGAAGTTGCTTTTTCAGCCAA	137917		
Qy	959	CTCATACTAATGTGCTGTGATGGCTTGCTTCTCAAGGAGTTTACTCAATTTTCACTGTG	1018		
Db	137916	AAATGAAAAATGTTTATCGGCGAAGTACTTTTCTAATGAAGGTTTCTCGCAACTCGCTATCG	137857		
Qy	1019	CGGAAACTGAGAAGTATCCCTCATGTTTCTTCTTTTAAATGTTGGTGGAGAACTTCAAT	1078		
Db	137856	CTGAAACAGAAAAATATCCACACGTAACNTTCTTTTGAATGGCGGAGAAATGAAGAAT	137797		
Qy	1079	TCCAAGATGAAGACGTTGTATGTTCCGTCACCAAAAGAGTTGCTTACATATGATTAA	1138		
Db	137796	TCCCTGTGAAACCCGATCTTAATCAATTCGCCAA--AAGTAGAAACATACGATTAA	137740		
Qy	1139	AACAGAAATGAATGCTGCTGGAGTTGCCGAAAATAATGGTCGACCAATTTGAGTCAGGCA	1198		
Db	137739	AACCTGAATGAGTCATATGAAGTAACAGATGCGCTTGTGTGAAGACATTTAAAACGACA	137680		
Qy	1199	GGCATCTTTGGTTATGTGCAATTTTGCCTCTCTGACATGTTGGCATGTTGGCATCTGTAAT	1258		
Db	137679	ACACGATGGATCATCTTAACCTTCGTAATCCAGATATGTTTGGACATTCAGGTATGC	137620		
Qy	1259	TTGAACCTGCCGTCACAAAGCATGTCAAGCTACTGACGAGGCAATTTGGAAGATATTTGAAG	1318		
Db	137619	TTGAGCCAACCTATTAAAGCAATCGAAGCAGTAGATGAAATCTTTGGTCTGTAGTAGATC	137560		
Qy	1319	CATGCCAACTTATAATTACGTTCTTATGTTTACTTCCGATCATGGAATGCTGAGAGA	1378		
Db	137559	TTATTTAGAAAAAGGTGGCTCAGCTATTATCTTTGTGACCACTGGTAACCTCTGAAACAA	137500		
Qy	1379	TGATTGCTCCGATGGTAGTGAACATACTGCACATACCTGCAATTTGTCCTCCATTTACT	1437		
Db	137499	TGCTACTCCAGAGGAAACCGCACACTGGCACACTACCGTTCCAGTTCACGTAAT	137441		
RESULT 6					
AX417048/c					
LOCUS	AX417048	349980 bp	DNA	linear	PAT 02-SEP-2002
DEFINITION	Sequence 4039 from Patent WO0228891.				
ACCESSION	AX417048				
VERSION	AX417048.1	GI:21449658			
KEYWORDS					
SOURCE	Listeria innocua				
ORGANISM	Listeria innocua				
REFERENCE	1				
AUTHORS	Kunst,F. and Glaser,P.				
TITLE	Listeria innocua, genome and applications				
JOURNAL	Patent: WO 0228891-A 4039 11-APR-2002;				
	INSTITUT PASTEUR (FR); CENTRE NATIONAL DE LA RECHERCHE				
	SCIENTIFIQUE (CNRS) (FR)				
FEATURES					
source	1. 349980				
	Location/Qualifiers				
	/organism="Listeria innocua"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:1642"				
	/note="seq 2058, original length: 3.011.208 replaced				
	by-seq 2058: 0.000.001 to 0.349.980-seq 4032: 0.300.001 to				
	0.649.980-seq 4033: 0.600.001 to 0.949.980-seq 4034:				
	0.900.001 to 1.249.980-seq 4035: 1.200.001 to				
	1.549.980-seq 4036: 1.500.001 to 1.849.980-seq 4037:				
	1.800.001 to 2.149.980-seq 4038: 2.100.001 to				
	2.449.980-seq 4039: 2.400.001 to 2.749.980-seq 4040:				
	2.700.001 to 3.049.980-seq 4041: 3.000.001 to 3.011.208"				
BASE COUNT	101702 a	72631 c	57863 g	117784 t	
ORIGIN					
Query Match	18.3%;	Score 314.2;	DB 6;	Length 349980;	
Best Local Similarity	53.3%;	Pred. No. 2.7e-58;			
Matches 767;	Conservative 0;	Mismatches 648;	Indels 24;	Gaps 4;	
Qy	14	AGTTTGAGTGGACAAATATCAAAATGTTCAACAAAAGTCTGTCTGTAGTTATTGATG	73		
Db	176453	ATTATTGGAGCAGTAATAATAATGAGTAATACCTGAGCAATATTATTCCTCGATG	176394		
Qy	74	GATGGGCGCTTTCGATGAAACACACCGGAATGCAATTTGCTTAAAGCTTAAACGCTATTA	133		

Db	176393	GTGTTGGTAAACGCTGCAGAACAGTAGTAATGCTGTAGCTCAAGCAAAACCAAAAT	176334
Qy	134	TGGACAAACCTTCTGTCGAAATGGCA---AAATTTGGAAGCACACGGTCTTCATGTTG	190
Db	176333	TCGACCGTTACTGGGCTAATTTTCCACGGGAACCTTAAAGCTGCTGGCTTGTATGTTG	176274
Qy	191	GATTGGCAGAGCTTAATGGGAAATTCGAAGTTGGACATTTGAATATAGGAGCTGGAA	250
Db	176273	GTCTTCAGAGGCCAAATGGTAACCTGGAAGTTGGCCATAACAACATCGGAGCTGGAC	176214
Qy	251	GAGTTATATCAAGATATTTGTCGAATTAATTTGGCTGTTCAACGAACGAGTTGTTA	310
Db	176213	GTATTTGTACCAAGCTTAACCTGATTTGATTAAGCAATTTGAAGAGCGCAATTCGAAG	176154
Qy	311	CAAACTCTCAGATTTGTGCATCAGCTGAGCGTCAAGAGAGGGAGTGGTCGATTCGATT	370
Db	176153	AGACAAAGCCCTAAACAATGCTTTTCACTCATACAAAGAAACAATTCGGACCTACATC	176094
Qy	371	TATTAGACTGGTTAGCGATGGTGGTCCACTCTCATATTTGATCATCTTTTTCGCTTGA	430
Db	176093	TTTTCGGTTTACTATCAGACGGCGGTGCATAGTCACATTAATCACCTCGTTGCACTTT	176034
Qy	431	TAGCTGCATTTAAACAATTAACAAGTCCAAAGTTTTCATCTACTTTTTCGTCGATGGTC	490
Db	176033	TAGAATCGCAAGATAAAGCGCTGAAAACGTTTATATCCATGCAATTCCTTGATGGAC	175974
Qy	491	GAGATACTTCGCCCAACAGTGGAGCTGGTTATCTTGAACAACCTCTTCAATTTATGCTT	550
Db	175973	GTACGTTGGCACCAATCTCTCACTAGATATTTTGAACACATAGAAAAGCTATTAGTG	175914
Qy	551	CGGAAAGTACGGAGAAATGGCTTACTTACTGACGTTTATTATGCAATGGATAGGACA	610
Db	175913	ACTTAACTATGTGCAATTTGCTACTGTTCTGACGTTTCTACGGATGGATCGTGATA	175854
Qy	611	AAAGATGGGAGCTATTAAAGATGGCTATGAGCAATTTGCGAGTATTTGACAAAG	670
Db	175853	AACGCTGGGAAGCTGTTGAAAAGCATACAAAGCATCGTGAAGCGTGAAGGTGAAAAA	175795
Qy	671	CCACCGTTGATAGGCTGTCGATGTTGTTAGAGAGCGATATGCTCAATCTGACACTGACG	730
Db	175794	-----TTTGAAGATCCAATCGAACTTGTCAAAGCTTCTTACGCTATGACAAAATGATG	175740
Qy	731	AATTTCTGAACCAATTTTTCGGACGATGGGC-----GAGTAAAGATG	778
Db	175739	AATTCGTTCTCTGCTATCATTTACTTAAAGATGGCAACCTGTTCAACAGATTAAGACA	175680
Qy	779	ACGATACTCTTATTTCTCAATTCGTCGTCATGCTATGCTCAAAATTTGTGAATGTT	838
Db	175679	ACGATCGATGATTTCTTCAATTTCCGTCGACGCGGATTCACCTTTCTAATGCAAT	175620
Qy	839	TGGGCTCGAAGCTTTAAGATCTTAAATAGTTCTGCTTCCCTCACCCCTAAAAATATTAGA	898
Db	175619	TCACTATAAAGAAATGGATCATTTTCGACGCTGGGCTAACCCACCTAAAAACATCAAT	175560
Qy	899	TTAGTCGGATGACCAATACAAATGAAGATTTCCATTTCCATGTTATTTCCACCTGGA	958
Db	175559	TCGTTAATGATGACTTTTCAACCCCAAGCATTTGCTGAAGTTGCTTTTGAGCCCAATG	175500
Qy	959	CTCATACTAATGCTGCTGATGCTGCTCTCAAGAGTTACTCAATTTCACTGTCG	1018
Db	175499	AAATGAAAATGTTATCGCGAAGTACTTTCTAATGAAGGTTTGTGCAACTGCGTATCG	175440
Qy	1019	CGGAACTGAGAAGTATCTCATGTTTACCTTCTTTAATGGTGGTCGAGAAGTTCAAT	1078
Db	175439	CTGAACAGAAAATATCCACAGTAACATTTCTTTATGAATGGCGACCAATGAAGAT	175380
Qy	1079	TCCAAGATGAAGCGTTGATGTTTCCGTCACCAAAAGATGCTACATATGATTTAA	1138
Db	175379	TCCCTGGTGAACCGGATTTCAATCAATTCGCCAA---AAGTAGAAACATACGATTTAA	175323
Qy	1139	AACCCAGAAATGATGCTGCTGGATGTCGCGAAAAAATGTCGAGCAAAATTTAGTCAGGCA	1198
Db	175322	AACCTGAATGATGATATGATGAAGTAACAGATGCGCTTGTGAAGCATTTAAACAGACA	175263
Qy	1199	GCATCCTTTGTTATGTGCAATTTTTCGCCCTCCTGCATGTTGGACATGTTGGTAAAT	1258
Db	175262	ACACGATCGGATCATCTTAAACTTCGCTAATCCAGATATGTTGGACATTCAGGTATGC	175203
Qy	1259	TTGAACCTCCGCTCAAAGCATGTCAGCTACTGACGAGGCAATTTGAAAGATATTTGAAG	1318
Db	175202	TTGAGCCAACTATTAAAGCAATCGAAGCAGTAGATGAAAATCTTGGCTGTAGTAGATC	175143
Qy	1319	CATGCCAACTTATTAATTAAGTTCTTATGTTTACTTCCCATCATGGAATGCTGAGAAGA	1378
Db	175142	TTATTTTGAAGAAAAGGTGGCTCAGCTATATCTTGTCTGACCATGTTAACTCTGAAACAA	175083
Qy	1379	TCATTTGCTCCGATGTTAGTGAACATGCTGACATACCTGCAATTTGTTGCCATTTACT	1437
Db	175082	TGCTACTCCAGAAAGAAAACCGCACACTGCGCACACTACCGTTCCAGTTCAGTAAT	175024
RESULT 7			
AL591983/C			
LOCUS	AL591983	324050 bp	DNA linear BCT 06-JUN-2002
DEFINITION	Listeria monocytogenes strain EGD, complete genome; segment 11/12.		
ACCESSION	AL591983	AL591824	
VERSION	AL591983.1	GI:16411809	
KEYWORDS	Listeria monocytogenes		
SOURCE	Listeria monocytogenes		
ORGANISM	Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.		
REFERENCE	1		
AUTHORS	Glaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Amend, A., Baquero, F., Berche, P., Bloeker, H., Brandt, P., Chakraborty, T., Charbit, A., Chetoui, F., Couve, E., de Daruvar, A., Dehoux, P., Domann, E., Dominguez-Bernal, G., Duchaud, E., Durant, L., Dussurget, O., Entian, K. D., Eshti, H., Portillo, F. G., Garrido, P., Gautier, L., Goebel, W., Gomez-Lopez, N., Hain, T., Hauf, J., Jackson, D., Jones, L. M., Kaerst, U., Kreft, J., Kuhn, M., Kunst, F., Kurapkak, G., Madueno, E., Maitournam, A., Vicente, J. M., Ng, E., Purcell, R., Remmel, B., Rose, M., Schlueder, T., Simoes, N., Tierrez, A., Vazquez-Boland, J. A., Voss, H., Wehland, J. and Cossart, P. Comparative genomics of Listeria species		
TITLE	Science 294 (5543), 849-852 (2001)		
JOURNAL	21537279		
MEDLINE	11679669		
PUBMED	2		
REFERENCE	(bases 1 to 324050)		
AUTHORS	Glaser, P., Frangeul, L. and Rusniok, C.		
TITLE	Direct Submission		
JOURNAL	Submitted (06-JUN-2001) Glaser P., Institut Pasteur, Genomique des Microorganismes Pathogenes, 25 rue du Docteur Roux, 75724 Paris Cedex 15, FRANCE		
COMMENT	E-mail: pglaser@pasteur.fr Phone: +33 1 45 68 89 96, Fax: +33 (0)1 45 68 87 86.		
FEATURES	Location/Qualifiers		
source	1..324050		
	/organism="Listeria monocytogenes"		
	/mol_type="genomic DNA"		
	/strain="EGD-e"		
	/db_xref="taxon:1639"		
gene	complement(290..492)		
	/gene="lmo321"		
CDS	complement(290..478)		
	/gene="lmo321"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAD00399.1"		
	/db_xref="GI:16411810"		
	/db_xref="SPTREMBL:Q928W7"		
	/translation="MAERIFRQTFNGSEIFIDDRTHIANPAFRQKIALIETGCEKMTDYTELKLAGYEEVTR"		
RBS	complement(487..492)		
	/gene="lmo321"		
gene	complement(586..801)		
	/gene="lmo322"		

```
CDS      complement(586. .801)
/gene="lmo2322"
/codon_start=1
/transl_table=11
/product="gp44 [Bacteriophage A118]"
/protein_id="CAD00400.1"
/db_xref="GI:16411811"
/db_xref="SPTREMBL:O8Y4V8"
/translacion="MSEVLVSASVEGYESKSNFTEINNIVKERFKIDEVERKRAE
TFNKKEVTKELVDGHLREIIPRTL"
gene      complement(join(798. .1331,1340. .1345))
/gene="lmo2323"
complement(798. .1331)
/gene="lmo2323"
/codon_start=1
/transl_table=11
/product="gp43 [Bacteriophage A118]"
/protein_id="CAD00401.1"
/db_xref="GI:16411812"
/db_xref="SPTREMBL:O8Y4V7"
/translacion="MFQKSTSATAAMQVLAETRTQKELADISVYTPALISNQIKGKRT
VSLEQABQLDSYNEPRSTYLFARHSNGMIPPLFDGLDNHSLTNRFEEVEEANN
TLKNGLTMTFNLKRGDMLQREAKQAISEITDVSALTLNASIAFAFNIDLOQVLN
KRDLYOKSGSLVSCGK"
gene      complement(810. .815)
/gene="lmo2322"
complement(810. .815)
/gene="lmo2322"
complement(1340. .1345)
/gene="lmo2323"
complement(1454. .2230)
/gene="lmo2324"
complement(1454. .2230)
/gene="lmo2324"
note="similar to anti-repressor [Bacteriophage A118]"
/codon_start=1
/transl_table=11
/protein_id="CAD00402.1"
/db_xref="GI:16411813"
/db_xref="SPTREMBL:O8Y4V6"
/translacion="MSNLQVIANMLPVLNEQGEKFDARMLEKILLVTRFNDWIY
RMIGNYENGLDFYSTLNGRPSNTNYFLTDTAKELAWQNNMGRAIRKIFIEV
EKQAKIATYPAPSYMIEDPVARAKKIEEQEKQVLKLEQKPKVVFARAVQTS
ENTILVLDATILKGLDIGNRFLFWLRGSGVLLSKGAYNKPQKAMNGLFGEQK
THIHDRNGLMKRTYTPQVTGKGQVLLNKLLEHNQVII"
gene      complement(2212. .2487)
/gene="lmo2325"
complement(2212. .2487)
/gene="lmo2325"
/codon_start=1
/transl_table=11
/protein_id="CAD00403.1"
/db_xref="GI:16411814"
/db_xref="SPTREMBL:O8Y4V5"
/translacion="MIFVNSFPQSGHVFTPGTLVNTCEHSGQLSPVFLIISCFAIF
RLCSLICLSNFESNLILPTLSLHKNTYTKRTERTKQIYK"
gene      complement(2493. .2774)
/gene="lmo2326"
complement(2493. .2774)
/gene="lmo2326"
note="similar to protein gp41 [Bacteriophage A118]"
/codon_start=1
/transl_table=11
/protein_id="CAD00404.1"
/db_xref="GI:16411815"
/db_xref="SPTREMBL:O8Y4V4"
/translacion="MNNIKAIKLETILENGNAIESGSFVKSVIKNLSLLEKDOE
LKTIEMKVLNGVEDSIENATPLLEKRLUEAKSLVEDLASTINSLEIKVK"
complement(2498. .2503)
gene      complement(2498. .2503)
/gene="lmo2325"
complement(2498. .2503)
/gene="lmo2325"
complement(2782. .2787)
/gene="lmo2326"
complement(2800. .3097)
/gene="lmo2327"
complement(2800. .3084)
/gene="lmo2327"
/transl_table=11
/protein_id="CAD00405.1"
/db_xref="GI:16411816"
/db_xref="SPTREMBL:O8Y4V3"
/translacion="MNRRLYSNKRSEQKDCVFKTDIPPENLKSINIKMOGDRNCCYG
VLEINGQLRKGTAVKLDLKAGSLPVVQVEYHPFTTISEMQRLWSGKY"
RBS      complement(3092. .3097)
/gene="lmo2327"
complement(3287. .3540)
/gene="lmo2328"
complement(3287. .3529)
/gene="lmo2328"
note="similar to transcription regulator"
/codon_start=1
/transl_table=11
/protein_id="CAD00406.1"
/db_xref="GI:16411817"
/db_xref="SPTREMBL:O8Y4V2"
/translacion="MRIWLKIREENGHLTQODVAELADVERTTYASTEQGRNPSVYN
AMRIASVLDFEWTFIFDDKVKREKSHFYLNEKQEARK"
RBS      complement(3535. .3540)
/gene="lmo2328"
3672. .4161
/gene="lmo2329"
3672. .3677
/gene="lmo2329"
3685. .4161
/gene="lmo2329"
note="similar to a putative repressor protein
[Bacteriophage A118]"
/transl_table=11
/protein_id="CAD00407.1"
/db_xref="GI:16411818"
/db_xref="SPTREMBL:O8Y4V1"
/translacion="MFGDILMKLRKNLTCSDIAKILGVARTTYSYEQNRNRPDAE
IQIRADIYDVSLDYLRGRKNNTADTIAAHIDNSAEEDEIKELATIEKREKRVNE
EEINITDAASKGDKVEVNEFDENEDFKAVARIMSDAEAVKAKVSFIEFYEQKNN"
4528. .5243
/gene="lmo2330"
4528. .4533
/gene="lmo2330"
4542. .5243
/gene="lmo2330"
note="similar to protein gp33 [Bacteriophage A118]"
/codon_start=1
/transl_table=11
/protein_id="CAD00408.1"
/db_xref="GI:16411819"
/db_xref="SPTREMBL:O8Y4V0"
/translacion="MSKYSYLLKKWFWMILLILLISLFGNGIWLFFATLATLTFAI
IKVKNNRRKYLIIATVLLTSLRTVOLYNYIVENPENVTIDNOKKDTKED
AETDEVEETEDQADASAAYVTQEGIESFESINRLISFNGILIKVVPFENEYDM
LIAYVTODLAKYQDETIIKKNADYLGNEIQRALGTFLGGDSNHRPVELRYEDETMA
GSSAFDKNMKLTGK"
5251. .5944
/gene="lmo2331"
5251. .5256
/gene="lmo2331"
5264. .5944
/gene="lmo2331"
note="weakly similar to gp32_Bacteriophage A118 protein"
/codon_start=1
```

/transl\_table=11

```
Query Match      18.2%; Score 312.6; DB 1; Length 324050;
Best Local Similarity 53.2%; Pred. No 6e-58;
Matches 766; Conservative 0; Mismatches 649; Indels 24; Gaps 4;

QY 14 AGTTTGAGATGGCAAAATATCAAAATGTTTCAACAAAAGTCTGCTTGTAGTTATTGATG 73
DB 14205 ATTATTGGAGCGAGTAAATAAATAAATGAGTAAATCACCTGTGACAAATTATTATCCTCGATG 134146

QY 74 GATGGGCGCTTTCGATGACACACACCGGAATGCTAAAGCTAAACCCCTATTATTA 133
DB 134145 GTTTTGTAAACGTGCAAGAACAGTAGGTAATCTGTAGCTCAAGCAACAAACCAAAAT 134086

QY 134 TGGACAAAACCTTTTCTCGAAATTTGGCA---AAAATTGGAAGCACACGGTCTTCATGTTG 190
DB 134085 TCACCGTATTGGCTGATTTCCTCACGGGAACCTTAAAGCTGTGCTTGTGATGTTG 134026

QY 191 GATTGCCAGAGCTTTAATGGGAATTTCTGAAGTTGGACATTTGAATATAGGAGCTGGAA 250
DB 134025 GCCTTCAGAAAGTCAAAATGGGTAACTCTGAAGTTGGCCATACAAACATCGGAGCTGGAC 133966

QY 251 GAGTTATTATCAAGATATGTCGAATTAATTTGGCTGTTCAACCAACAGGATTTGTTA 310
DB 133965 GTATTGCTACCAAAAGCTTAACCTGATTGACAAAGCTATTGGAAGGCGGAATTCCTCAAG 133906

QY 311 CAAATCCTCAGATTGTTGCATCAGCTGAGCGTCAAGAAAGGGAGCTGGTGCATTCGATT 370
DB 133905 AGAACAAAGCCCTAAACAATGCTTTTCACATACAAAGAAACAACTCGGACTTACATC 133846

QY 371 TATTAGAGCTGGTTAGCGATGGGTGTCACCTCTCATATGATCAATCTTTTTCGCTTGA 430
DB 133845 TTTTTCGGTTTACTATCAGACGGCGGTGTCATAGTCACATTAATCACCTCGTTGCATTT 133786

QY 431 TAGCTGCATTTAAACAATTAACAAGTCCCAAGTCTTTCATCTACATTTTTCGTCGATGTC 490
DB 133785 TAGAACTCGGAAGATAAGCGGTGAAAACGCTTTATATCCATGCTTCTTGTGAGGAC 133726

QY 491 GAGATATCTCGCAACAAGTGGAGCTGGTATCTTGAACAACCTCTTCAATTTATTGCTT 550
DB 133725 GTGACGTGCGCACCACAACTCTCCTACTAGATAATTTAGAAACACTACAAAAAGCTATTAGT 133666

QY 551 CGGAAAGTACGGAGATTTGGCTACTATTACTGAGCGTTATTATGCAATGATAGGAC 610
DB 133665 ACTTAACATATGCGCGGATTTGCTGTTCTGAGCGTTTCTACGCGATGGATCGTGATA 133606

QY 611 AAAGATGGGAGCTATTAAAGATGGCTTATGAGGCAATTTGAGGATTTGGACAAAG 670
DB 133605 AACGCTGGGAACGTGTTGAAAAGCAATACAAAGCAATCTGACGCTGAAGTGAAAAA- 133547

QY 671 CCACCGTTGATAAGGCTGTCGATGTTGTTAGAGAGCGATATGCTCAATCTGAGACTGACG 730
DB 133546 -----TTTGAAGATCCAATCGAACTGTGCAAAAGCTTCTTATGCTATGACAAAAATGATG 133492

QY 731 AATTTCTGAACCAATGTTTTTTCGGAGCATGGGC-----GAGTAAAGATG 778
DB 133491 AATTGCTTCTGCTATCATTTACTATAAGATGGCAAACTGTGTCAACAGTTAAAGACA 133432

QY 779 AGGATCTCTTATTTCTCAATATATGCTGCTGATGCTGATGCTGCTGCTGCTGCTGCTGCT 838
DB 133431 ACGATGCACTTATTTCTCAATTTCCGCTCTGACGCTGCGATTAACCTTTCTTAATGCAAT 133372

QY 839 TGGGTCTCGAAGCTTATAAGATCTTAATAGTTTCGTTCTCCACCTCAAAAAATATTTCAGA 898
DB 133371 TCACTGATAAAGATGGGATCATTTTCGACCGTGGGCGAGATCACCTCAAAAAACATCAAT 133312

QY 899 TTAGTGGGATGACCAATACATAAAGATTTCCATTTCCATCGTTATTTCACCTGTA 958
DB 133311 TTGTTACATAGTACTCTTTACAAACCAAGATTTGATGCTGAAGTTGCTTTTGGAGCAATG 133252

QY 959 CTCATCTAATGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1018
DB 133251 AATGCAAAATGTTATCGCGAAGTACTTTCTAATGAAGGCTTATCGCAACTGCGTATCG 133192
```

```
QY 1019 CGGAAACTGAGAAGTATCCTCATGTTTACCTTCTTCTTAATGCTGCTGAGAAATTCAT 1078
DB 133191 CTGAACACAGAAATATCCACAGTAACATTTCTTTATGATGCTGACGAAATGAAGAT 133132

QY 1079 TCCAAGATGAAGCGCTTGTATGTTTCCGTACCAAAAAGATGTTGCTACATATGATTTAA 1138
DB 133131 TCCTGTTGAAAACCGGATTTCTAATCAATTCGCAA--AAGTAGAAAACATGCAATTTGC 133075

QY 1139 ACCAGAAATGATGCTGCTGGAGTTGCCGAAAAAATGTCGAGCAAAATGATCGAGGCA 1198
DB 133074 AACCTGAAATGAGCGCATATGAAGTAACAGATGCACTTGTGTAAGACATTTAAAAAGCA 133015

QY 1199 GGCATCCTTTGGTTATGTGCAATTTTGGCCTCTCCTGACATGTTGGACATACCTGATAAT 1258
DB 133014 AACACGATCGATCATCTTAACTTCGAAACCCAGATATGTTGGGCACTCAGGTATGC 132955

QY 1259 TTGAACCTGCCGTCAAAAGCATGTCAGCTACTGACGAGGCAATGTTGAAAGATATTTGAAG 1318
DB 132954 TTGAGCAACGATTAAGGCAATCGAAGCAGTAGATCAAAATCTTGGTGTGTAGTAGACC 132895

QY 1319 CATGCCAACTATATATTAGTTCTTATGTTTACTTCCGATCATGAAATGCTGAGNAGA 1378
DB 132894 TTATTTTGAAGAAAGTGGTTGAGCTATTATCTTCTGCTGACCATGTAATCTGGAACAA 132835

QY 1379 TGATTGCTCCGATGCTAGTGAACATPACTGCACATACCTGCAATTTGGTCCCATTTACT 1437
DB 132834 TGCTACTCCAGAAAGAAACCCGACACTGCGCACACTCCGTTCCAGTCCAGTAAT 132776
```

RESULT 8  
AX641672/c

```
LOCUS      AX641672      349980 bp      DNA      linear      PAT 21-FEB-2003
DEFINITION Sequence 2862 from Patent WO0101118.
ACCESSION  AX641672
VERSION     AX641672.1  GI:28474433
KEYWORDS
SOURCE      Listeria monocytogenes
            Listeria monocytogenes
ORGANISM    Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
REFERENCE   1
AUTHORS     Buchrieser, C., Frangeul, L., Couve, E., Rusniok, C., Psihi, H.,
            Dehoux, P., Dussurget, O., Chetoui, F., Nedjari, R., Glaser, P.,
            Kunat, F., Cossart, P., Daniels, J., Goebel, W., Kreft, J., Kuhn, M.,
            Ng, E., Vasquez-Boland, J., Dominguez-Bernal, G., Garrido-Garcia, P.,
            Tierrez-Martinez, A., Amend, A., Chakraborty, T., Domann, E., Hain, T.,
            Berche, P., Charbit, A., Durant, L., Perez-Diaz, J. C., Baquero, F.,
            Garcia del Portillo, F., Gomez-Lopez, N., Madiueno, E., de Pablos, B.,
            Wehland, J., Kaerst, U., Entian, K. D., Hauf, J., Rose, M. and Voss, H.
TITLE       Listeria monocytogenes genome, polypeptides and uses
JOURNAL     Patent: WO 0101118-A 2862 11-APR-2001;
            INSTITUT PASTEUR (FR)
```

FEATURES  
Location/Qualifiers

```
source      1..349980
            /organism="Listeria monocytogenes"
            /mol_type="genomic DNA"
            /db_xref="taxon:1639"
            /note="Seq 1 for big: original length: 2944528-split
            into 10 <23>seq 0001 : From 0.000.001 to 0.349.980
            length: 349.980 <23>seq 2855 : From 0.300.001 to
            0.649.980 length: 349.980 <23>seq 2856 : From 0.600.001
            to 0.949.980 length: 349.980 <23>seq 2857 : From
            0.900.001 to 1.249.980 length: 349.980 <23>seq 2858 :
            From 1.200.001 to 1.549.980 length: 349.980 <23>seq 2859
            : From 1.500.001 to 1.849.980 length: 349.980 <23>seq
            2860 : From 1.800.001 to 2.149.980 length: 349.980
            <23>seq 2861 : From 2.100.001 to 2.449.980 length:
            349.980 <23>seq 2862 : From 2.400.001 to 2.749.980
            length: 349.980 <23>seq 2863 : From 2.700.001 to
            2.944.528 length: 244.528"
BASE COUNT  101075 a 74760 c 61320 g 112825 t
ORIGIN
```

Query/Match	18.2%	Score 312.6	DB 6	Length 349980
Best Local Similarity	53.2%	Pred. No. 6e-58		
Matches	766	Conservative 0	Mismatches 649	Indels 24
Gaps	4			
QY	14	AGTTTGAGATGGCAAAATATCAAAATCTCAACAAAAAGTCTGCTCTCTAGCTTATTGATG	73	
DB	129205	ATTATTGGAGCAGTAATAATAAATGAGTAANTCACCTGTAGCAATTAATTCTCTCGATG	129146	
QY	74	GATGGGCGCTTTCGGATGAACAACACGGGAATGCAATTTGCTTAAAGCTAAAAACGGCTATT	133	
DB	129145	GTTTGTCTAAACGTGCAGAACACAGTAGTAATGCTGTAGCTCAAGCAACAAACCAAAATT	129086	
QY	134	TGGACAACCTTTGTTCTGGAAATTTGGCA--AAATTTGGAAGACACGGTCTTCATGTTG	190	
DB	129085	TCGACCGTTTATTTGGGCTGATTTTCTCACGGGAACTTAAAGCTGCTGGCCTTGATGTTG	129026	
QY	191	GATTGCCAGAGGCTTAATGGAAATTTCTGAAGTTTGACATTTGCAATATAGAGCTGAA	250	
DB	129025	GCCTTCAGAGGTCAAAATGGTAACCTCTGAAGTTGGCCATACAAACATCGAGCTGGAC	128966	
QY	251	GAGTTAATTATCAAGATATTGTTCCGAATTAATTTTGGCTGTTTCAACGAACAGAGTTGTTA	310	
DB	128965	GTAATTGCTACCAAAAGCTTAACTCGTATTGACAAAGCTATTGAAGAGCGCAATTC	128906	
QY	311	CAAACTCTCAGATTGTTGCAATCAGCTGAGCGTGCAAAGAGGGAGTGGTCGATTCG	370	
DB	128905	AGAACAAAGCCCTAAACAATGCTTTCACCTACATACAAAGAAAAACCTCGGACTTACATC	128846	
QY	371	TATTAGACTGGTTAGCGATGGTGGTCCACTCTCATATTGATCATCTTTTTCGGTTGA	430	
DB	128845	TTTTTCGGTTTACTATCAGAGCGGCGGTGCATGCTACATTAATCACTCGTTCGCAC	128786	
QY	431	TACGTGCATTTAAACAATTACAAGTGCACAAAGGTTTTCATTCACCTTTTTCGTGATG	490	
DB	128785	TAGAACTGCGAAGATTAAGCGCTGAAACGTTTATATCCATGTCATCTCTGATGAC	128726	
QY	491	GAGATACTTCGCCAACAGTGGAGCTGGTTATCTTTGAACAACCTTCTTCAATTTATG	550	
DB	128725	GTGACGTGGCACCACAACTCTCACTAGAAATATTTAGAAACACTACAAAAAGCTATT	128666	
QY	551	CGGAAAAGTACGGAGAATTGGCTACTATTACTGACGCTTATATGCAATGATAGGACA	610	
DB	128665	ACTTAACATATGGCGGAGTTCCTACTGTTCTTCGACGTTTCTACGCGATGATCGTGA	128606	
QY	611	AAAGATGGGAGCGATTAAAGATGCTTATGAGCAAAATTTGTTGGAGGTTATGGCAAAA	670	
DB	128605	AACCTCGGAACGTTTGAAAGACATACAAAGCAATCGTAAGCGCTGAAGGTGAAAA	128547	
QY	671	CCACCGTTGTAAGCGTCTCGATGTTGTTAGAGACGATATGCTCAATCTGAGACTGAC	730	
DB	128546	-----TTTGAAGATCCAATCACTGTTCAAAAGCTTCTTATGCTAATGACAAAAATG	128492	
QY	731	AAATTCTGAACCAATTTTTCGGACGATGGC-----GAGTAAAGATG	778	
DB	128491	AATTGCTGTTTCCTGCTATCAATTAAGATGCGAAACCTGTTGCAACAGTTAAAGACA	128432	
QY	779	ACGATACTCTTATTTCTCAATTTATCGTCTGATCGTATGCGTCAAAATTTTGTAATG	838	
DB	128431	ACGATGAGTTATTTCTCAATTTCCGCTCTGACCGTGCATCACTTTCTAATGCTAT	128372	
QY	839	TGGGTCTCGAACGTTATAAGATCTTAATAGTTCGGTTCTCACCCCTAAAAATATCAGA	898	
DB	128371	TCACGTATAAAGATGGGATCATTTTCGACCGTGGGGCAGATCACCCCTAAAAACATCA	128312	
QY	899	TTAGTGGGATGACCCCAATACAAATAAGAGTTTCCATTTCCATCTGTTATTTCCACCTG	958	
DB	128311	TTGTTCAATGACCTTTTACAAACCAAGTATTGATGCTGTAAGTTGCTTTTGAGCCAA	128252	
QY	959	CTCATACTAATGCTCTGCTCAATGGGCTTCTCTCAAGGAGTTACTCAATTTTCACTGTG	1018	
DB	128251	AAATGAAAAATGTTATCGCGAAGTACATTTCTTAATGAAGGCTTATCGCAACTGCGTAT	128192	
QY	1019	CGGAACTGAGAAGTATCCTCATGTTACCTTCTTCTTTAATGGTGGTGCAGAAAGTTCA	1078	

```

RESULT 9
LOCUS      AP003190/c
DEFINITION Clostridium perfringens str. 13 DNA, complete genome, section 6/10.
ACCESSION  AP003190 BA000016
VERSION     AP003190.2 GI:18146728
KEYWORDS
SOURCE      Clostridium perfringens str. 13
ORGANISM    Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium.
REFERENCE   1
AUTHORS     Shimizu,T., Ohtani,K., Hirakawa,H., Ohshima,K., Yamashita,A., Shiba,T., Ogasawara,N., Hattori,M., Kuhara,S. and Hayashi,H.
TITLE       Complete genome sequence of Clostridium perfringens, an anaerobic flesh-eater
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 99 (2), 996-1001 (2002)
MEDLINE     21664373
PUBMED      11792842
REFERENCE   2 (bases 1 to 301150)
AUTHORS     Shimizu,T.
TITLE       Direct Submission
JOURNAL     Submitted (15-FEB-2001) Tohru Shimizu, Institute of Basic Medical Sciences, University of Tsukuba, Department of Microbiology; 1-1-1 Tennohdai, Tsukuba, Ibaraki 305-8575, Japan
E-MAIL      tshimizu@md.tsukuba.ac.jp, Tel:81-298-53-3354, Fax:81-298-53-3354)
COMMENT     On Jan 14, 2002 this sequence version replaced gi:18144942.
FEATURES
             Location/Qualifiers
             1..301150
                /organism="Clostridium perfringens str. 13"
                /mol_type="genomic DNA"
                /strain="13"
                /db_xref="taxon:195102"
                /note="anaerobic pathogen for gas gangrene"
                complement(3)..5066)
                /gene="CPE1281"
                complement(3)..5066)
                /gene="CPE1281"
                /note="1687 aa, partially similar to pir:T36472 probable secreted alpha-galactosidase from Streptomyces coelicolor (680 aa); 33.6% identity in 140 aa overlap. Putative N-terminal signal sequence was found by PSORT"
                /codon_start=1
                /transl_table=1

```







/strain="QM B1551"  
/db\_xref="taxon:1404"  
34..1569  
/gene="pgm"  
34..1569  
/gene="pgm"  
/codon\_start=1  
/transl\_table=11  
/product="2,3-bisphosphoglycerate-independent  
phosphoglycerate mutase"  
/protein\_id="AAB26327.1"  
/db\_xref="GI:4589056"  
/translation="MSKKPVALIILDFALRDEKGNVTHAKPNFDRFWEYPHAT  
LQAGSAGVLPEQMGSEVGHNLIGAGRIYQSLFRVNVIREGFEQNETLLAAVK  
HAKKGTNLHLLGGLSDGVHSHIYALLRAKSEGLEKIVYIHGFLDGRDVAQSA  
ETYLKELNIEYGVGEIATLSGRVYSNDROKRWERKSVYRWYGGPSYSAEE  
CYKDSYDNGIYDFEVLPSVITKEDGSPVATIQDEDAVIFYNRPDRAIQISNIFANED  
FRSDFRGEKHPFLVCLTHFSETDGVYAFKPNLDNLTGLVLSNNLKQIRIAT  
EKYPHTVFMGSGREAFPGETRIILDSPKVATYDLKPEMSAYEVDLALAEIGDKQ  
DAILNFPAMDVGHSGMLEPTVKAIETVDECLGKIVDAILAKGGTATITADHGNAD  
VITLEGNPMATHNTVPVIVTKQGLREDEGLDGLAPMLTLFLDVAOPKEMTKT  
LIK"

BASE COUNT 540 a 318 c 363 g 433 t  
ORIGIN

Query Match 17.4%; Score 299.6; DB 1; Length 1654;  
Best Local Similarity 52.4%; Pred. No. 9.8e-55;  
Matches 737; Conservative 0; Mismatches 654; Indels 15; Gaps 3;

QY 42 TCAACAAAAGTCTGCTGTAGTATTGATGATGGGCGCTTCCGATGGAACAACAGG 101  
DB 39 TAAAGAGCCATAGCATTAATCATCTAGATGCTTTTGCATTACGGCATGAAGATAAG 98  
QY 102 GAATGCAATTGCTAAAGCTAAACGCTTATTATGAGCAAACTTTGCTCGAAATGGCA 161  
DB 99 TAATGCGGTAAACATGCAAAAACCAAACTTCGACCGTTTCTGGAACGAGTATCCTCA 158  
QY 162 ---AAAATGGAAAGCACAGCGTCTTCATGTTGGATTGCCAGAGGCTTAAATGGGAAATTC 218  
DB 159 TGCTACATCTACGCGCTCTGGAGAAGCTGTAGGCTTACCAGAGGCCAAATGGGTAAC 218  
QY 219 TGAAGTTGGACATTTGAATATAGGAGCTGGAAGAGTATTATCAAGATATTGTTGCAAT 278  
DB 219 TGAAGTAGGTCATTAACATCGCGCAGCGCGCAATTTGTATACCAAGCTTAACTCGTG 278  
QY 279 TAATTTGGCTGTTCACAGAAACAGTGTGTTACAAATCTCAGATTGTTGCATCAGCTGA 338  
DB 279 AAACGTAGCAATTCGTGAAGCGAGTTTGAACAAACGAAACGCTTCTAGCAGCTGAAA 338  
QY 339 GCGTGAAGAAGAGGAGTGGTCGATTGCAATTTATTAGGACTGGTTAGCGATGGTGGTGT 398  
DB 339 GCACGCAAAAGAAAGGATACAAACCTTCATCTTTTCGGGCTTTTATCTGATGGCGGTGT 398  
QY 399 CCACTCTCATATTGATCACTTTTTCGTTGATACGTGCTATTAAACAATTACAAGTGC 458  
DB 399 ACACAGTCATATGAACATCTATATGCAATTTGCGCTAGCTAGAGTGAAGCTTAGA 458  
QY 459 AAAGGTTTCACTCACTTTTTCGTCATGTCGAGATACCTTCGCCAACAAAGTGGAGCTGG 518  
DB 459 AAAGTCTACATTCATGCTTCTTATAGTGTGCTGACGTAGCACCTCAATCTGCTGAAC 518  
QY 519 TTATCTTGAAACACTTCTTCAATTTATTTGCTGGAAAAAGTAGCGAGAAATTTGGCTACTAT 578  
DB 519 GTATTTAAAGAACTTAATGAAGAAATTTGAAGAGTACGGCGTTGGTGAATTTGCAACACT 578  
QY 579 TACTGACGCTTATTATGCAATGATAGGACAAAGATGGGCGCTATTAAAGTGGCTTA 638  
DB 579 TCAGAGCGTACTACTCAATGACCGCAGCAAAACGCTGGGCGGTAGAGAAATTCATA 638  
QY 639 TGAGGCAATTTGAGGAGTATTGGACAAAAAGCCACCGTTGTATAGGCTGTGCTATCTGT 698  
DB 639 TCGCGCAATGGTCTACGCGCAAGGTTCATCATATACAAAGCGCTGAAGAGTGGCTAAAGA 698

QY 699 TAGAGAGCGATATGCTCAATCTGAGACTGACGAAT-----TTCTGAAACCAATGT 749  
DB 699 TTGCTATGACAATGGAATCTATACGAATTCGTTTACCTTCTGTTATACAAAGAAGA 758  
QY 750 TTTTTCGGACGATGGCGAGTAAAGATGACGACTATCTTATTTTCTTGAATTCGTGC 809  
DB 759 CGGATCGCGAGTTGCGACAATTCAGCAAGAGATGAGTATTTTCTACAACTTCGCTCC 818  
QY 810 TGATCTGATGCGTCAAAATTTTGTGAATGTTTGGTCTCGAAGCTTATAAGATCTTAAATAG 869  
DB 819 GGATCGCGAATTCAAATTTTCAATAGCTTCGCAACGAAGATTTCCGTTTCATTTGATCG 878  
QY 870 TTCGTTCTCTCAACCTTAAATAATATCAGATTAGTGGGATGACCAATACAAATAAGAGTT 929  
DB 879 CGGTGAAAAACATCCAAAACACTTACATTTTGTATGTTTAACTCACTTCACGCGAAACAGT 938  
QY 930 TCCATTTCCATCGTTTATTTCCACCTGTGACTCATATGCTTGTCTGCTGAAATGGCTTGC 989  
DB 939 TGACGGATATGATAGCCCTTTAAAGCCAATTAACTTGGATAACACGCTTGGTGAAGTATATC 998  
QY 990 TTCTCAAGGAGTTTACTCAATTTTCACTGTGCGGAAACTGAGAAGTATCCTCATCTTACCTT 1049  
DB 999 TCNAATAACTTAAACAGCTTCGTATTGCGGAGACAGAAAATATCCTCACGTAAGCTT 1058  
QY 1050 CTCTTTAATGTTGGTGCAGAAATTTCAATTTCCAAAGTGAAGCGCTTGTATGTTCCGTC 1109  
DB 1059 CTTTATGCGGTGGAGCTGAAGCAGAGTTTCTGCTGGAACACAGTATCTTAATCGATTC 1118  
QY 1110 ACCAAAGAAGTTCCTACATATGATTTAAACCAGAAATGAATGCTGCTGAGTTCGCGA 1169  
DB 1119 ACCAA---AAGTAGCAACATATGACTTGAACCTGAGATGAGTGTATGAAGTACGGA 1175  
QY 1170 AAAATGTCGAGCAAAATTTGAGTCAGGAGGAGTCTTTGTTTATGTCGAAATTTTCGCGC 1229  
DB 1176 CGGTTGCTGCAAAATCGAAGCGATAGCAGAGCTATTTTATTAACCTTTCGAAA 1235  
QY 1230 TCGTACATGTTGGACATCTGTAATTTTGAACCTGCGCTCAAAAGCATGTCAAGCTAC 1289  
DB 1236 TCGTATGTTGGAGCATTTCAGTATGTTAGAACCACCTGTAAAAGCGATTCGAAACAGT 1295  
QY 1290 TGACGAGGCAATTTGGAAGATATTTCAAGCATGCCAACTTATAATACGTTCTTATGTT 1349  
DB 1296 AGATGAGTCTTAGGCAAAATTTGATAGCAATTTTAGCTTAAAGCGGTACACAACTCAT 1355  
QY 1350 TACTTCGATGTAATGCTGAGAAGATGATGCTCCGATGCTGAGTGAACATCTGTC 1409  
DB 1356 CACAGCAGCACTGATGATGCTGAGTGAAGTATACGCTTGAAGGTAATTCATGACGCG 1415  
QY 1410 ACATACCTGCAATTTGGTCCCAATTA 1435  
DB 1416 TCATACACGAATCTGTTCCAGTAA 1441

# RESULT 11

AE017014/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

AE017014 300854 bp DNA linear BCT 16-MAY-2003  
Bacillus cereus ATCC 14579 section 17 of 18 of the complete genome.  
AE017014 AE016877  
AE017014.1 GI:29898525  
Bacillus cereus ATCC 14579  
Bacillus cereus ATCC 14579  
Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus; Bacillus  
cereus group.  
1 (bases 1 to 300854)  
Ivanova,N., Sorokin,A., Anderson,I., Galleron,N., Candelon,B.,  
Kapatral,V., Bhattacharya,A., Reznik,G., Mikhailova,N.,  
Lapidus,A., Chu,L., Mazur,M., Goltsman,E., Larsen,N., D'Souza,M.,  
Walunas,T., Grechkin,Y., Pusch,G., Haselkorn,R., Fonstein,M.,  
Ehrlich,D.S.D., Overbeek,R. and Kyrpides,N.  
Genome sequence of Bacillus cereus and comparative analysis with  
Bacillus anthracis  
Nature 423 (6935), 87-91 (2003)





OTWYIAKTLNLEASGVVFAVSVPRLEIFMIVGAVADRFPPKKNIMFYNGIIRAILVA  
TILTWIVGDITLYTALFALFFGLADAFWSADGSLPVLVEKSLRTOANSITOMTN  
QASVILGVGLILKFTNVEITISITILLVAAILVOKIOFTPEOKDOKWETS  
KEGILYVESPLSLFCSAFLNLFILGPHQVGPPLFVKVNLHGDLSOFYSLEASV  
GGMAIGAVVGLKNRRRGLFCIIMLLSGVFFLSINFSTVLMQALLAGMFYGITI  
ANAIPLMAMIOSIVKEEMGRVMSLLMLSSMGFIPLSYAFSTIALAIGIPIVTWKS  
GAIAIVFVFAIRVPVVRKFD"  
complement(15379..16266)  
/locus\_tag="BC4930"  
complement(15379..16266)  
/locus\_tag="BC4930"  
/codon\_start=1  
/transl\_table=11  
/product="Transcriptional regulator, Deor family"  
/protein\_id="AAP11803.1"  
/db\_xref="GI:29898531"  
/translation="MGSLYNLAKPYSEFOSKEEFNYQKVLKHYRFQNLKNDIIRH  
FGKYAVNEKQTVGACPLMETIATNVKRSINTVRSIAKLEELGIIIRKATKRRH  
RGYISANLYVFLISAIDRMDRMKSAESEDYTAGCSKEQKSERETILSKNIPIQIK  
EKRTIYLEDYTRHDIPKPIFYALLPMTSNPKINIFWSKVELAYKKSGLLEQVL  
LEQILADEEYVGNFIRWRSVVRVAYKGEIRKNVRALLYSTVRDLFLEVLEWGAVR  
RSKGIPLFPFKKEPCVA"  
16347..16808  
/locus\_tag="BC4931"  
16347..16808  
/locus\_tag="BC4931"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical Cytosolic Protein"  
/protein\_id="AAP11804.1"  
/db\_xref="GI:29898532"  
/translation="MIVISACLGACRYDGNLNVSKTEELLOKEDTVLICEVLGG  
LETPRPSABIIIGDNDVDLDGKAVMTKDEVDTEAFVSGYKALEQIKNLHPEYIIL  
KERSPSCGSSTIYGFENGNKOTGFGVTTALPKRHKFKVISEEENKRN"  
complement(16839..17192)  
/locus\_tag="BC4932"

Query Match 17.4%; Score 299; DB 1; Length 300854;  
Best Local Similarity 52.6%; Pred. No. 5.6e-55;  
Matches 728; Conservative 0; Mismatches 640; Indels 15; Gaps 3;  
QY 65 TTATTGATGGGCGCTTCGGATGAACACACACGGGAATCAATGCTAAAGCTAA 124  
DB 205559 TTCCTGACGCTTCGGACCTCTGTAAGAACTTACGGGAATGCTGTAGCAGCTAAGA 205500  
QY 125 CGCCTATTATGACAAACTTTGCTCGAAATTTGCA---AAATTTGGAAGCACACGGCTC 181  
DB 205499 AACCTAAATTTGATGTTTGAACAAATTTCCCTCACACAACGCTTACAGCTTGTGGC 205440  
QY 182 TTCATGTTGGATGCCAGAGCTTAATGGGAATTCGAAGTTGGACATTTGAATATAG 241  
DB 205439 AGGCAGTAGGCTTCCAGAAAGTCAAAATGGGTAACCTGAGGTTGGTCACTTAAATATCG 205380  
QY 242 GAGCTGGAAGAGCTATTATCAAGATATTGTTGGAATTAATTTGGCTGTTCACAGAAAGC 301  
DB 205379 GTGCTGGCGCATTTGATATCAAGCTTAACACGGTTAACCTTGCAATTCGTGAAGGTG 205320  
QY 302 AGTTGTTTACAATCCTCAGATGTTGTGATCAGCTGAGCGTCAAGAGAGGGAGGTGGTC 361  
DB 205319 AGTTGATAGAAGAAAGCTTCCAAAGTCAATTAAGCGCTAAAGAAAAAGGTACTG 205260  
QY 362 GATTGCATTTATAGGACTGGTTAGCGATGGTGGTGTCCACTCTCATATTGATCATCTTT 421  
DB 205259 CACTTCATTTATTCGGTTTACTTTCTGACGGGTTGTCACAGTGCACATGAACCATATGT 205200  
QY 422 TTGCGTTGATACCTGCATTTAAACAAATTAACAAGTGCACAAAGCTTTTTCATCTACTTTTG 481  
DB 205199 TTGCTCTTCTCGCTTAGCAGCTAAAGAGGGCTTGAGAAAGTATACATTCATGATCT 205140  
QY 482 CTGATGGTGGAGATCTTCGCCAACAAAGTGGAGCTGGTTATCTTGAACAACTTCTCAAT 541  
DB 205139 TAGATGGCGCGCATGTTGGACCAAAAACAGCAAAAGCTATATCGATGCAACAAATGAAG 205080

QY 542 TTATTGCTTCGGAAAAGTACGGAGAAATGGCTACTATTACTGACGCTTATTATGCAATGG 601  
DB 205079 TAATTAAGAAACAGAGTAGACAATTCGGACATATCTGCTGCTTATTACTCCATGG 205020  
QY 602 ATAGGACAAAAGATGGGAGCGTATTAAAGATGGCTTTATGAGGCAATGTTGGAGATTTG 661  
DB 205019 ACCGTGACAAGCGTTGGATCGCTAGAAAAATGTTACCGTGTCTATGGTGAATGGTGAAG 204960  
QY 662 GACAAAAACCCACCGTTGATAAGGCTGTCGATGCTTCTTATAGAGAGCGATATGCTCAATCTG 721  
DB 204959 GGCCTACTTATAATCAGCAGAGAGAGTGTGTAAGACTCTTATGCGAICGGTATCTATG 204900  
QY 722 AGACTGACGAATTTCTGAAACCAATTTGTTTTTTCGGACGATGGCGAG-----TAA 772  
DB 204899 ATGAATTCGATTATTCGCTCTGTAATTTTAAACGAAGATAACACGCCAGTTGCAACAATCA 204840  
QY 773 AGATGACACATCTCTTATTTCTTCAATATATGCTGCTGCTGATGCTGCTGCTCAAAATTTG 832  
DB 204839 ATGATGATGATGACGATTATCTTCTATAACTTCGCTCCAGACCGTGAATTAATTAATGCTC 204780  
QY 833 AATGTTGGGCTCTCGAAGCGTTATAAAGATCTTAAATAGTTTCGGTTCCTCACCTTAAAAATA 892  
DB 204779 GTGATTTACAACGAAGACTTCCGCTGAGTTCGATCGTGGTGAAGAAAGTACCTCACATTC 204720  
QY 893 TTCAGATTAGTGGGATGACCCCAATACAATAAAGAGTTTCCATTTCCATCTGTTATTCACAC 952  
DB 204719 CAGAAATTCGATGATGACACACTTCAGTGAAACTGTAGATGTTACGTTGCGATTCACAG 204660  
QY 953 CTGTCACATCACTAATGTTGCTGCTGTAATGCTGCTGCTTCTCAAGAGATTACTCAATTTTC 1012  
DB 204659 CAATGAACCTTGAATCACTTTTAGGTGAAGTTTGTGCACAAGCGGATTAAGACCAACTTC 204600  
QY 1013 ACTGTGCGGAAACTGAGAAAGTATCTCATGTTTACCTTCTTCTTAAATGTTGTCGAGAAG 1072  
DB 204599 GCATCGCGGAAACTGAAGAGTATCCGACGTTACATTTCTTCTTTAGCCGTTGCTGAGG 204540  
QY 1073 TTCAATTCAGATGAAGAGCGTTGATGGTTCCGTCACCAAAAGAAAGTTGCTACATATG 1132  
DB 204539 CTGAATTTCCAGGAGAAAGAGCGTATCTTAATTAACCTCACCAA---AGGTTGCAACGTATG 204483  
QY 1133 ATTTAAACCAAGAAATGATGCTGCTGGAGTTGCCGAAAAAATGTCGAGCAAAATGAGT 1192  
DB 204482 ACTTGAACCTGAAATGAGCATTTACGAAGTAAACGACGCTTTAGTAAATGAATCGAAA 204423  
QY 1193 CAGGACGAGCTCCTTTGGTTATGTGCAATTTTTCGCGCTCCTGACATGGTTGACATACCTG 1252  
DB 204422 ATGATAAATCATGATGTTATCTTAACTTTTGGCAACTGTGATATGTTGGCCATTCG 204363  
QY 1253 GTAAATTTGAACCTCCGTCACAGCATGTCAGCTACTGACGAGGCAATTTGGAAGATAT 1312  
DB 204362 GGATGATGGAACCAACAATTAAGCAGTAGAAGCAACTGACGAATGTTTAGAAAAAGTTG 204303  
QY 1313 TTGAAGCATGCCAAACTTATAATTTAGTTTCTTATGTTACTTCCGATCATGAAATGCTG 1372  
DB 204302 TAGAAGCATTTCTTCCAAAAGATGGTGTAGCAGCTTATTACTGCTGACCATGTAATGCGG 204243  
QY 1373 AGAAGATGATTGCTCCCGATGAGTGAACATPACTGCACATACCTGCAATTTTGGTCCCAT 1432  
DB 204242 ATGAAGAATTAACCTTCTGAAGGAGAGCAAACTGACAGCTCATACAACCTAACCCGTTCC 204183  
QY 1433 TTA 1435  
DB 204182 TCA 204180  
RESULT 12  
AE017040/c  
LOCUS AE017040  
DEFINITION Bacillus anthracis str. Ames section 17 of 18 of the complete genome.  
ACCESSION AE017040  
VERSION AE017040.1  
KEYWORDS GI:30259630

AE017040  
Bacillus anthracis str. Ames section 17 of 18 of the complete genome.  
AE017040 AE016879  
AE017040.1 GI:30259630

SOURCE	Bacillus anthracis str. Ames
ORGANISM	Bacillus anthracis str. Ames
	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE	1 (bases 1 to 291205)
AUTHORS	Read, T., Peterson, S., Tourasse, N., Baillie, L., Paulsen, I., Nelson, K., Tettelin, H., Fouts, D., Eisen, J., Gill, S., Holtzapple, E., Oksstad, O., Helgason, E., Rilstone, J., Wu, M., Kolonay, J., Beanan, M., Dodson, R., Brinkac, L., Gwinn, M., DeBoy, R., Madupu, R., Daugherty, S., Durkin, A., Haft, D., Nelson, W., Peterson, J., Pop, M., Khouri, H., Radune, D., Benton, J., Mahamoud, Y., Jiang, L., Hance, I., Weidman, J., Berry, K., Plaut, R., Wolf, A., Watkins, K., Nierman, W., Hazen, A., Cline, R., Redmond, C., Thwaite, J., White, O., Salzberg, S., Thomason, B., Friedlander, A., Koehler, T., Hanna, P., Kolsto, A.-B. and Fraser, C.
TITLE	The genome sequence of Bacillus anthracis Ames and comparison to closely related bacteria
JOURNAL	Nature 423 (6935), 81-86 (2003)
MEDLINE	22608414
PUBMED	12721629
REFERENCE	2 (bases 1 to 291205)
AUTHORS	Read, T., Peterson, S., Tourasse, N., Baillie, L., Paulsen, I., Nelson, K., Tettelin, H., Fouts, D., Eisen, J., Gill, S., Holtzapple, E., Oksstad, O., Helgason, E., Rilstone, J., Wu, M., Kolonay, J., Beanan, M., Dodson, R., Brinkac, L., Gwinn, M., DeBoy, R., Madupu, R., Daugherty, S., Durkin, A., Haft, D., Nelson, W., Peterson, J., Pop, M., Khouri, H., Radune, D., Benton, J., Mahamoud, Y., Jiang, L., Hance, I., Weidman, J., Berry, K., Plaut, R., Wolf, A., Watkins, K., Nierman, W., Hazen, A., Cline, R., Redmond, C., Thwaite, J., White, O., Salzberg, S., Thomason, B., Friedlander, A., Koehler, T., Hanna, P., Kolsto, A.-B. and Fraser, C.
TITLE	Direct Submission
JOURNAL	Submitted (26-MAR-2003) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
FEATURES	Location/Qualifiers
source	join(1..170086,186846..291205)
	/organism="Bacillus anthracis str. Ames"
	/mol_type="genomic DNA"
	/strain="Ames"
	/db_xref="taxon:198094"
	/focus
	complement(170087..186845)
source	/organism="Bacillus anthracis phage lambda Ba02"
	/proviral
	/mol_type="genomic DNA"
	/db_xref="taxon:229344"
	/note="Putative prophage; Putative lysogenic prophage region. Best Hits not completed. Putative attL/R"
gene	90..803
	/locus_tag="BA5145"
CDS	90..803
	/locus_tag="BA5145"
	/note="similar to GP:4584089, and GP:4584089; identified by sequence similarity; putative"
	/codon_start=1
	/transl_table=11
	/product="conserved hypothetical protein"
	/protein_id="AAP28817.1"
	/db_xref="GI:30259631"
	/translation="MHFVRLAQDHEFTANONPKAEPMARYMKNHPFGIGIQTPERRQ LKDKVIQHTLPDKDQIIVRELDWPERFQAALDMQYKKKHINETHIFPEEL IVTKSWSDIVPTIFGNIFLHPELISAYIPKWTASDNIWLRQAAILFOLKYRQK MDELLFWVIGQLHSSKEFFIQKAIQWLVRYAKTKPDVVVYQNNELAPLSKREAI KHKENYGITNEKIGETLS"
gene	974..1105
	/locus_tag="BA5146"
CDS	974..1105
	/locus_tag="BA5146"
	/note="identified by Glimmer2; putative"
	/codon_start=1
	/transl_table=11
	/product="hypothetical protein"
	/protein_id="AAP28818.1"
	/db_xref="GI:30259632"
	/translation="MDAPFYQLGWRNQTLVPRQLSLGMSVFLCFSKHILLVSVQT 1151..2527
	/gene="glys"
	/locus_tag="BA5147"
CDS	1151..2527
	/gene="glys"
	/locus_tag="BA5147"
	/note="similar to SP:P75425; identified by sequence similarity; putative"
	/codon_start=1
	/transl_table=11
	/product="glycyl-tRNA synthetase"
	/protein_id="AAP28819.1"
	/db_xref="GI:30259633"
	/translation="MYSMEQVYVNLAKHGRFVPGSEYVGGTANTWYDGLGILKNNVLL KKAWKFIQESPVNGLDAAILNPKMTWIASGVGNFNDPMIDCKKAKRHADKLL EDALDAKIENVDGLTFDQMDLDMKEHYKCPDCGSEETEIROENLAKFPQGTVE SSTNEIFLRPETAGQIFVFNKNVQRSMRKLPGFIGOIGKSFREITPGNFTFRREF EQMELEFFCKPEDLEWFAFWRETCRNWLLSLGMEISRLDRHBEELSHSYNATTD IEFKFPFGWELGWASRTDFDLKRHRHNEDEFNFTDPQTNERYVTEIYPSLGRADR TFLAFDQABEEOLENDSRTVLRHFPALAPYKAAIPLPSKLSEGATEVFAELAKDF MVDFDEGTSICKRYRRODEICTPCTCYTDFEDSDVEDKAVTVYDRDRTWQVRMPISLKL FLEKKIQE"
gene	complement(2565..2944)
	/locus_tag="BA5148"
CDS	complement(2565..2944)
	/locus_tag="BA5148"
	/note="similar to GP:10119873; identified by sequence similarity; putative"
	/codon_start=1
	/transl_table=11
	/product="comA operon protein, putative"
	/protein_id="AAP28820.1"
	/db_xref="GI:30259634"
	/translation="MAKTLMDSLGIELLEMTEDKVATMPVDERTHOPFGFLHGGASL ALAETVASVSYNLDIQEKICFGLGIEINANHIRSKDKGMVTAIGTPIHRGHSTMVMDV RIIDENDLLCISRCTVAIKREKREK"
gene	3044..3787
CDS	3044..3787
	/locus_tag="BA5149"
	/locus_tag="BA5149"
	/note="similar to SP:P19196, GB:X53368, GB:248169, PID:48574, and PID:793895; identified by sequence similarity; putative"
	/codon_start=1
	/transl_table=11
	/product="conserved domain protein"
	/protein_id="AAP28821.1"
	/db_xref="GI:30259635"
	/translation="MKOKFFTLSLMFSLTLTPNCLAAKASNLITDVQVTQKQKKP YIEQVIRNPSPHNSDSKFNKLFYVKSTDQPKTRLEKDAKKYKKEGTSPTPRP YVANDVSVLNKPKPPLSLYLNYQYTGGAHGLYTWKANTFDLKEKKLLHDLDFQGE DKYEVIRAEIRKIKONEGIYFPDATEKVMSTKKFFHFFLEPNNLVYIFPLYEIAIYPS SGIPQFRIPYTLRLREYLKPSYQNILDNK"
gene	3838..4431
CDS	3838..4431
	/locus_tag="BA5150"
	/locus_tag="BA5150"
	/note="similar to GP:10175981, and SP:P04028; identified by sequence similarity; putative"
	/codon_start=1
	/transl_table=11
	/product="bio family protein"
	/protein_id="AAP28822.1"
	/db_xref="GI:30259636"
	/translation="MRRFHVLDLALAAMFVALMAIGANIVSWAPFLQVAGIPLSMOPFF FAILAGLLGSRGLASMTVYLVGVAGAPIFANFKAGFAGLLDPTGCTIIAFTIYV VSGKIVEQKEKPTSTFAISFTGIIITYIIGTTYMYGAVNLFMGNGNSYKAAWIMIM WFAVKDVIYFTIIGALIAPRIYYAVRSAYQHSHTIS"
gene	4579..5415
CDS	4579..5415
	/locus_tag="BA5151"
	/locus_tag="BA5151"

```
/locus_tag="BA5151"
/notes="identified by match to PFAM protein family HMM
PF02978"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP28823.1"
/db_xref="GI:30259637"
/transl_table=11
/translation="MNLKNTLKVSVGLVSLFFILGMSIGYKLVNSELGMSYTFGN
GLAFLIITIVSLCAALFIVIGLIKVKLPKAKSVATSIILFVTSIISLIVLFT
ITKVNMEEXYQVLOKQKKKEASLYLVAASFYNNINFNFAASYLVSESTWNSNAD
KQDFNALUSSRTEDIGMITVDYFISTMGNDLKLVSAAKEQPNKYKETEEYKKI
YGIILNFOAPSGLSIFNONVNLIOEYKKAAGNINIAITDEIKSNELKPTD
KN"
gene complement(5521..6408)
/genes="gtab"
/locus_tag="BA5152"
complement(5521..6408)
/genes="gtab"
/locus_tag="BA5152"
/notes="identified by match to TIGR protein family HMM
TIGR00453"
/codon_start=1
/transl_table=11
/product="Urp-glucose-1-phosphate uridylyltransferase"
/protein_id="AAP28824.1"
/db_xref="GI:30259638"
/transl_table=11
/translation="MKRVKAIIPAGLGTFLPATKAMPKEMLPIDVKPTIQYIVVEE
AVKSGIEDIIIVTGTKRSIEDHFNAPLEQNLEKKYELLEKVOASSKVDHYI
RKEPKGLGHAVWCARKFIGDEPFAVLGDDIVQAEKPCRLQRIEYDLKLSVIGVQ
TYPEDTHRGYIIDPLEQGRYQVRNFVFAAGTAPSLAIMRYLYLTPFIEMFLE
OQHVAGSGEIQIDTAIQSLNEIQRFVAYDFEGKRYDVGEKLGIVQTTIEMALQHPER
DMVVTMMKKILIEQTNQES"
gene complement(6517..8241)
/locus_tag="BA5153"
complement(6517..8241)
/locus_tag="BA5153"
/notes="identified by match to PFAM protein family HMM
PF02878"
/codon_start=1
/transl_table=11
/product="phosphoglucomutase/phosphomannomutase family
protein"
/protein_id="AAP28825.1"
/db_xref="GI:30259639"
/translation="MNWQEFSRWLSYABDLKLEQLENMKQDEKKIBDSFYKNLEF
GTGMRGELAGTNLNVYTVRKATGLAKFTIEKLEAEAKRGVVVAYVDSRHKSPFA
MEVAATLGAHGITTIVFESLRPTPLVLSFAVRLHIVSGIVLTASHNPPYNYKYVGE
Query Match 17.0%; Score 292.6; DB 1; Length 291205;
Best Local Similarity 52.3%; Pred. No. 1.4e-53;
Matches 724; Conservative 0; Mismatches 644; Indels 15; Gaps 3;
QY 65 TTATTGATGATGGGCCCTTCCTCGATGACACACACCGGAATGCAATGCTAAAGCTAAA 124
DB 189570 TTCTTGACGGCTTCGTGGAAGAACTTACGGGAATGCTGTAGCAAGCTAAGA 189511
QY 125 CGCCTATTATGACAAACTTTCTCTGGAATTTGGCA---AAAATTTGGAAGCACACGGT 181
DB 189510 AACCTAAATTTGATGGTTACTGGAACAAATTCCTCCACACACGCTTACAGCTTGTGGCG 189451
QY 182 TTCATGTTGGATGCCAGAGCTTAATGGGAAATTCGAAATTTGAGATATATAG 241
DB 189450 AGGCAGTAGGCTTCCAGAGGTCAAATGGGCAACTCTGAGGTGTGCTAAATATCG 189391
QY 242 GAGCTGGAAGACTTATTATCAAGATATGTTGCAATTAATTTGGCTGTTCAACGAACG 301
DB 189390 GTGCTGGCGGCTGTATATCAAGCTTAACACGGGTAAACGTTGCAATTCGTGAAGGTG 189331
QY 302 AGTTTGTACAAATCCTCAGATGTTGTCATGAGCTGAGCGTGCAAGAGAGGGAGTGGTC 361
DB 189330 AGTTTCGAAGAACGACGATTTCCAAAGTGCATTAAGCGCTAAAGAAAGAGGTACTG 189271
```

```
QY 362 GATTGCAATTTATTAGGACTGGTTAGCGATGGTGGTGTCCACTCTCATATTTGATCATCTTT 421
DB 189270 CACTTCATTTATTGGTGTCTTCTGACGGTGGTGTGCACAGTGCACATGAACACATGT 189211
QY 422 TTGCGTGTGATAGCTGCATTTAAACAATTAACAAGTGCCAAAGGTTTTTCATTCACCTTTT 481
DB 189210 TTGCTCTTCTTCCGCTTAGCAGCTAAAGAAGGCGTTGAGAAGGTATACATTCATCATCT 189151
QY 482 CTGATGGTCGAGACTCTTCCGCCAACAAAGTGGAGCTGTTTATCTTTGTAACAACTTCTTCAAT 541
DB 189150 TAGATGGCGCGCATGTGTGACCAAAAACAGCACAAAGCTATATCGATGCAACAAATGAAG 189091
QY 542 TTATTGCTTCGGAAAAGTAGTCGGAGAATTTGGCTACTATTACTGGACGCTTATTATGCAATGG 601
DB 189090 TAAATAAAGAAACAGGAGTAGGACAATTCGCGACTATCTCTGCTGCTTATTACTCCATGG 189031
QY 602 ATAGGGACAAAAGATGGGAGCGTATTAAAGATGCTTATGAGCAATTTGTTGGAGGTATTG 661
DB 189030 ACCGTGACAAAGCGTTGGGATCGGTAGAAAATGTTTACCGTGTATGGTGAATGGTGAAG 188971
QY 662 GACAAAAGCCACCGTTGATAAGGCTGTGATGCTTGTAGAGAGCGATATGCTCAATCTG 721
DB 188970 GCCTTACTTATAATCAGCAGAGAGAGTGTAGAACACTTATGCAATGGTATCTACG 188911
QY 722 AGACTGACGAATTTCTGAAACCAATTTGTTTTTCGGACGATGGCGAG:-----TAA 772
DB 188910 ATGAATTCGTATTGCGCTCTGTAATTTTAAACAAGATAACACGCCAGTTGCAACAATCA 188851
QY 773 AGATGACACACTCTTCTTATTTCTTCAATATTCGTGCTGATCGTATGCGTCAAAATTTG 832
DB 188850 ATGATGATGATGACGTATCTTCTATAACTTCGCTCCAGACCGTGCATTTCAAAATTTG 188791
QY 833 AATGTTTGGTCTCGAACGTTTATAAGATCTTAATAGTTTCGGTTCCTCACCTCAAAAATA 892
DB 188790 GTGATTTACAACAGGAGACTTCGCTGAGTTTGCATCGTGGTGAAGAAAGTACCTCACATTC 188731
QY 893 TTCAGATTTAGTGGGATGACCCCAATACAATAAAGAGTTTCCATTTCCATTCGTTATTCCTCC 952
DB 188730 CAGAAATTCGATGATGACACACTTTAGTGAACACAGTGGATGGTTCGTTGCAATCAAG 188671
QY 953 CTGTGACTCATACTAATGCTTGTGCTGAATGGCTGCTTCTCAAGGAGTTACTCAATTTTC 1012
DB 188670 CAATGAACCTTGATACACTTTTAGGTGAAGTTGTTCCGCAAGCGGATTTAAAGCAACTTC 188611
QY 1013 ACTGTGCGGAAACTGAGAAAGTATCCCTCATCTGTACCTTCTTCTTTAATGTTGGTGCAGAA 1072
DB 188610 GCATCGGAAACTGAAATAATATCCGACGTTTACATTTCTTTTACGCGTGGTGCAGG 188551
QY 1073 TTCAATTCAGATGAAGAGCGTTGATGTTCCGTCGTCACCAAAAGAGTTGCTACATATG 1132
DB 188550 CTGAATTCGAGGAGAGCGTATCTTAAATTAACCTCACCGA---AGTTGCAACGATG 188494
QY 1133 ATTTAAACCAAAATGAATGCTGCTGGAGTTGCGGAAATTTGCGGAAATTTGCTGAGCAATTTG 1192
DB 188493 ACTTGAACCTGAAATGAGCATTTAGGAAGTAAACGAGCGCTTTAGTAAATGAATCGAA 188434
QY 1193 CAGGAGGACTCCTTTGTTGTTATGCAATTTTCCGCTTCTGACATGGTTGACACTACTG 1252
DB 188433 ATGATAAATCATGATGTTATCTTCTTAACTTTCCGAACTGTGATATGTTGGCCATCTG 188374
QY 1253 GTAATTTGAACCTGCGCTCAAGCATGTCAGCTACTGACGAGGCAATTTGAAAGATAT 1312
DB 188373 GGATGATGAAACCAACAATTAAGCAGTAGAAGCAACTGACGAACTTTAGGAAAGTTG 188314
QY 1313 TTGAAGCATGCCAACTTATATACGTTCTTTATGTTACTTCCGATGCTGAAATGCTG 1372
DB 188313 TAGAAGCGATTTGCAAAAGATGTTGATAGCTTATTACTGCTGACCATGTTATGCTG 188254
QY 1373 AGAAGATGATTTCTCCCGATGTTAGTGAACACTACTGACATACCTGCAATTTTGGTCCCAT 1432
DB 188253 ATGAAGAAATTAATTTCTGAAGGAGAACCAATCACAGCTCATACACTAACCGGTTCC 188194
QY 1433 TTA 1435
```









Db 6917 GAGATACAGATGGGAAAGAGAAGAACTTGCCTTACAATGCAATGGTCTTGGAAAAAGGTG 6976  
Qy 665 AAAAAAGCCAGCTGTGATAGAGCTGTGATGCTTGTAGAGAGCGATATGCTCAATCTGAGA 724  
Db 6977 AGAGGCGAAGAGTGTATTAAGCTGTAGATGCTTCATATCATGACACAAACAGATG 7036  
Qy 725 CTGACGAATTTGAAACCAATGTTTTCGG-----ACGATGGCGGAGTAAAGATG 778  
Db 7037 AATTGTACTACCAACAGTATATAGTTAAAGAAGGAAACAGGTTGCGACTATAAGAGACA 7096  
Qy 779 ACATATCTTATTTCTCAATATATCGTGTATGCTGATGCTGAATTTGTAATGTT 838  
Db 7097 AGGATCTGTATATTTCTCAATATTTAGCTGTATGCTGATGCTGATGCTGATGCTGAT 7156  
Qy 839 TGGGTCTCGAACCTTATAAGATCTTAATAGTTGCTGCTCACCCTCAAAATATTGAGA 898  
Db 7157 TACAGAGAGCTTTTGCGGATTTAAGAGAGACAGACTTAATATGAGTTGTTGTAAC 7216  
Qy 899 TTAGTGGGATGACCCCAATACAAATTAAGAGTTTCCATTTCCATGTTTATCCCACTGTGA 958  
Db 7217 TGACGAATATGATGCAAGCTTTAAGGAGTTGATGCTG-----ATTGGACCTGAAA 7270  
Qy 959 CTCATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1018  
Db 7271 ACATCAAAATCTCTTGGAGATATGTAAGCAACAAAGGTTTAAATCACTTAGAATG 7330  
Qy 1019 CGAAATGAGAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1078  
Db 7331 CAGAACTGAAATATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7390  
Qy 1079 TCCAAGATGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1138  
Db 7391 ATAAAAATGAAGATAGAGCTTTAATATATCATCAACAA---AGTTGCAACTTATGATCT 7447  
Qy 1139 AACCAAAATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1198  
Db 7448 AACCAAAATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7507  
Qy 1199 GGCATCTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1258  
Db 7508 AGTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7567  
Qy 1259 TTGAACCTGCGCTCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1318  
Db 7568 TTGAACGAGCTAAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7627  
Qy 1319 CATGCCAACTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1378  
Db 7628 AGTCTTAAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7687  
Qy 1379 TGATTCCTCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1410  
Db 7688 TGATCATCTACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT

## RESULT 15

U39725/c  
LOCUS 12844 bp DNA linear BCT 19-OCT-2000  
DEFINITION Mycoplasma genitalium section 47 of 51 of the complete genome.  
ACCESSION U39725 U39730 L43967  
VERSION U39725.1 GI:3845017  
KEYWORDS  
SOURCE Mycoplasma genitalium  
ORGANISM Mycoplasma genitalium  
REFERENCE 1 (bases 1 to 12844)  
AUTHORS Fraser,C.M., Gocayne,J.D., White,O., Adams,M.D., Clayton,R.A., Fleischmann,R.D., Bult,C.J., Kerlavage,A.R., Sutton,G.G., Kelley,J.M., Fritchman,J.L., Weidman,J.F., Small,K.V., Sandusky,M., Fuhrmann,J.L., Nguyen,D.T., Utterback,T., Saudek,D.M., Phillips,C.A., Merrick,J.M., Tomb,J., Dougherty,B.A., Bott,K.F., Hu,P.C., Lucier,T.S., Peterson,S.N., Smith,H.O. and Venter,J.C.  
TITLE The minimal gene complement of Mycoplasma genitalium

## JOURNAL MEDLINE PUBMED REFERENCE AUTHORS

Science 270 (5235), 397-403 (1995)  
96026346  
7569993

2 (bases 1 to 12844)

Fraser,C.M., Gocayne,J.D., White,O., Adams,M.D., Clayton,R.A., Fleischmann,R.D., Bult,C.J., Kerlavage,A.R., Sutton,G.G., Kelley,J.M., Fritchman,J.L., Weidman,J.F., Small,K.V., Sandusky,M., Fuhrmann,J.L., Nguyen,D.T., Utterback,T., Saudek,D.M., Phillips,C.A., Merrick,J.M., Tomb,J., Dougherty,B.A., Bott,K.F., Hu,P.C., Lucier,T.S., Peterson,S.N., Smith,H.O., Hutchinson,C.A. III and Venter,J.C.

## Direct Submission

Submitted (29-OCT-1995) The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20850, USA

3 (bases 1 to 12844)

Fraser,C.M., Gocayne,J.D., White,O., Adams,M.D., Clayton,R.A., Fleischmann,R.D., Bult,C.J., Kerlavage,A.R., Sutton,G.G., Kelley,J.M., Fritchman,J.L., Weidman,J.F., Small,K.V., Sandusky,M., Fuhrmann,J.L., Nguyen,D.T., Utterback,T., Saudek,D.M., Phillips,C.A., Merrick,J.M., Tomb,J., Dougherty,B.A., Bott,K.F., Hu,P.C., Lucier,T.S., Peterson,S.N., Smith,H.O. and Venter,J.C.

## Direct Submission

Submitted (19-OCT-1998) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA

On or before Oct 19, 2000 this sequence version replaced

gi:1046149, gi:1046106.

## FEATURES source

1..12844

/organism="Mycoplasma genitalium"

/mol\_type="genomic DNA"

/isolate="G37"

/db\_xref="taxon:2097"

125..1474

/gene="MG425"

125..1474

/gene="MG425"

/note="similar to GB:U00089 SP:P75172 PID:1673885 percent identity: 65.56; identified by sequence similarity;

putative"

/codon\_start=1

/transl\_table=4

/product="ATP-dependent RNA helicase (dead)"

/protein\_id="AAC72446.1"

/db\_xref="GI:3845018"

/translation="MDSTFHELGISOTLIETLNLHINKPTKIQOISIPQFLSEKNLI  
VHSPGCTKTAFAPIPIELKLEKEDQAKTIVAPTRVLPQIKTFSNIAKKNKL  
IISLGGVPKQIKKIKTNPOIIVGTGRIIMDLERKAIHFSDEHLHIDEVDLMD  
RGFKQIFNLLEQINSFKQIAVSASYNQEAIAIKQITNNGIFIGSPFENKADNTNN  
DKLIQKQFYLFSDQKQALYSLIKTAQVKSIIIVFCDTKLVDDHVFURKNELRTFA  
LHGDKQFIRERNLAFANTKOPTILVTDLIGRGHVEADIMVINYSAELKATYIN  
RMGRTRNNHKGTCVTCTSOEKKVFLKWEKIDTNRKIAECKOMKIKLPLKNKATK  
KGIISLDCVQKIYANAKPYDRNKRPVSLASDLFKSRMRQPEKAKQKIHNDWQSNM"  
1497..1691

/gene="MG426"

1497..1691

/gene="MG426"

/note="similar to GB:U00089 SP:P75171 PID:1673884 percent identity: 84.38; identified by sequence similarity;

putative"

/codon\_start=1

/transl\_table=4

/product="ribosomal protein L28 (rpl28)"

/protein\_id="AAC72447.1"

/db\_xref="GI:3845019"

/translation="MAKKDQLTGLRGLYGNRSHSKTITRRKWNVNLOSCKIKDTNGK  
VTRILVSTKNTITLKKQNR"  
1727..2152

/gene="MG427"

1727..2152

/gene="MG427"

/note="similar to GB:U00089 SP:P75170 PID:1673883 percent identity: 75.89; identified by sequence similarity;

putative"





Db 5608 ATGCAAGGTAGTATTACATTTATTTGGTGATGGTAGAGATGTAGCACCTTGTAGCTTAA 5549  
 QY 515 CTGGTTATCTTGAACAACTTCTCAATTTATTGCTTCGGAAGTAGCGGAGATTCGCTA 574  
 Db 5548 ACAAGATCTTGGAGAAATTAAGATATTTCTAAAAAATATCCTAATGTGTATTGGAA 5489  
 QY 575 CTATTACTGGACGTTATTATGCAATGAGGAGGACAAAGATGGGAGCGGTATTAAAGATGG 634  
 Db 5488 CTATTGGGAGAGATGATGAGATGATGATGATGATGATGATGATGATGATGATGATG 5429  
 QY 635 CTATGAGGCAATTTGTTGAGGT-----ATTGACAAAAAGCCACCGTTGATAAGGCTG 688  
 Db 5428 CTATAAAGCTTTATTAGGAGTTTCAAAAAATTAATCAATGACCAATTTGGTTATATTG 5369  
 QY 689 TCGATGTTCTAGAGAGCGATGTCATCTGAGACTGACGAAATTTCTGAAACCAATTG 748  
 Db 5368 AAAGCAATATCAGAACCAATTTACTGATGAATTTATTTCTGCAATTAATGCCAATT 5309  
 QY 749 TTTTTCGGACGATGGCGAGTAAAGATGAGGATGACTCTTATTTTCTTCAATTTATCGTG 808  
 Db 5308 TAAATTTCTGATCAGTTTGCATTAACAATAATGATGGAGTTATTTCCCTTTAATTTAGAC 5249  
 QY 809 CTGATCGTATGCTCAAAATTTGTAATTTGGGCTCGAACGTTATAAAGATCTTAATA 868  
 Db 5248 CTGATAGCAAGACAAATGTCCTATTTGATCTTTAACAGCAATTTATTAACAATCAAC 5189  
 QY 869 GTTCGGTCTCCACCTTAAATAATTTACAGATTAGTGGGATGACCCCAATACATAAGACT 928  
 Db 5188 CTGAATTTGAACGAAAGAAATTTATTTTGTAAATGATGAATTTATGAGGGAATTG 5129  
 QY 929 TTCCATTTCCATCGTTTATTTCCACCTGTGACTACTACTTATGCTTGTGCTGAATGGCTTG 988  
 Db 5128 TACCTAGCGAATTTGCTTTTCCACCTCAAAACCATTAATAATAGTCTTGTGAGTAATTG 5069  
 QY 989 CTCTCAAGGATTTACTCAATTTACTGTGCGGAACCTGAGAGTATCCTCATGTTACT 1048  
 Db 5068 CTAATAATAATTTAGCAATTTGAGGATTTGAGAACTGAAAGATGCTCAGTTACTT 5009  
 QY 1049 TCTTCTTTAATGTTGTCGAGAGTTCAATTTCAAGATGAGGCGTTCTATGTTCCGT 1108  
 Db 5008 TCTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 4949  
 QY 1109 CACCAAGAAGTTGCTACATATGATTTAAACCAAGAAATGATGCTGCTGGAGTTGCCG 1168  
 Db 4948 CTTAA---AAGTGTACATATGATTTAGTCCCGAAATGTCATGTAAGCTATTACTG 4892  
 QY 1169 AAAAAATGCTCAGCAAAATTTGAGTCAAGCAGGATCCTTTGGTTATGTCGAATTTGGCC 1228  
 Db 4891 ATGCACACTAGAAAAGCTTAA-----TAACCTTTGATTTTACTGTTTAAATTTTGCTA 4838  
 QY 1229 CTCTGACATGTTGGACATCTGTTAAATTTGAACCTGCCGTCAGCATGCTCAAGCTA 1288  
 Db 4837 ATCTGATATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATG 4778  
 QY 1289 CTGACGAGCAATTTGGAAGATTTTGAAGCATGCGCAAACTTATTAATACGTTCTTATGG 1348  
 Db 4777 TCGATGTTCAAAATTAACCAATAGTTGATTTTGTAAAGCTAATCAATAAATGATGTTT 4718  
 QY 1349 TTACTTCCGATCATGGAATGCTGAGAGATGATGCTCCCGATGCTGAGTGAACATCTG 1408  
 Db 4717 TAACTGAGATCATGGGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 4658  
 QY 1409 CACATACCTGCAATTTGGTCCCATTTACTTCTCTCCAAACATTTGTTTAAATCGA 1468  
 Db 4657 ACACACATTTAATCTGTACCATTTGATGTTGATGTTGATGTTGATGTTGATGTTGATG 4598  
 QY 1469 CT 1470  
 Db 4597 CT 4596